

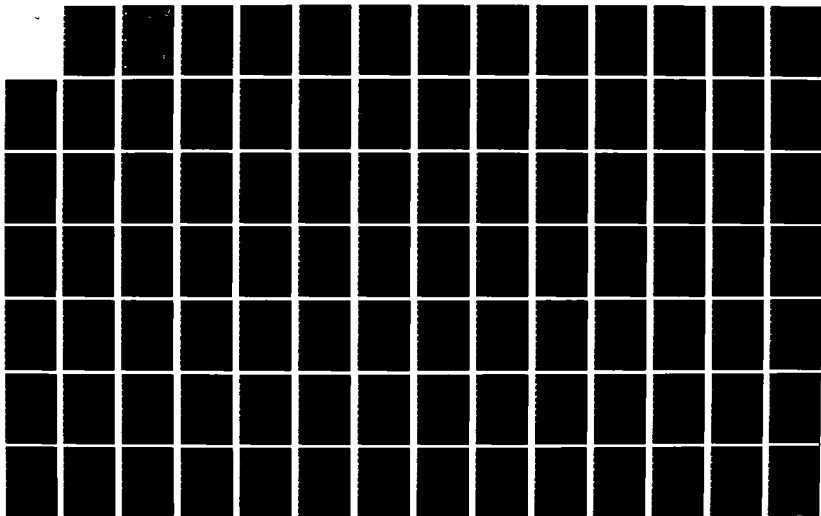
AD-A138 029

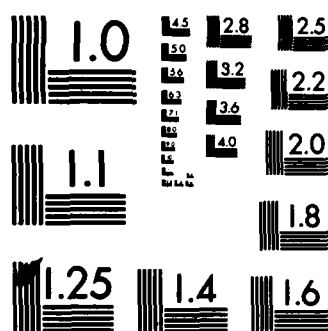
EXPLORATION OF DOSE-RESPONSE TECHNIQUES WITH SOME
APPLICATIONS TO A SIMUL. (U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH SCHOOL OF ENGI... L G KEHL
DEC 83 AFIT/MA/GOR/83D-4 F/G 9/2

1/2

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A138029



EXPLORATION OF DOSE-RESPONSE TECHNIQUES
WITH SOME APPLICATIONS TO A
SIMULATION PROBLEM

THESIS

Larry G. Kehl
First Lieutenant, USAF

AFIT/MA/GOR/83D-4

DTIC
S ELECTE
FEB 21 1984
D

DTIC FILE COPY

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

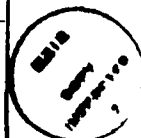
DISTRIBUTION STATEMENT A

Approved for public release;
Distribution is unlimited.

84 02 17 073

AFIT/MA/GOR/83D-4

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A/1	



EXPLORATION OF DOSE-RESPONSE TECHNIQUES
WITH SOME APPLICATIONS TO A
SIMULATION PROBLEM

THESIS

Larry G. Kehl
First Lieutenant, USAF

AFIT/MA/GOR/83D-4

DTIC
ELECTE
S FEB 21 1984

D

Approved for public release; distribution unlimited

AFIT/MA/GOR/83D-4

EXPLORATION OF DOSE-RESPONSE TECHNIQUES WITH SOME
APPLICATIONS TO A SIMULATION PROBLEM

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Operations Research

Larry G. Kehl, B.S.
First Lieutenant, USAF

December 1983

Approved for public release; distribution unlimited

Preface

This report proposes to investigate and apply various quantal response models to determine the applicability of these methods on data generated by the Avionics Evaluation Program. It is hoped that this effort will be a basis for moving quantal response methods out of the realm of bio-assay and into more general applications.

I wish to acknowledge my indebtedness to my thesis committee, Dr. Joseph Cain (reader), and to the ever patient Dr. David Barr, whose suggestions and guidance were invaluable to this effort. I would like to give a special thanks to Dr. Thomas W. Copenhaver, of Wyeth Laboratories, for his support of this effort, and particularly for the computer source code and program deck which he gave me. I would also like to thank my wife, Li Hua, for her support and understanding during this effort.

Larry G. Kehl

Table of Contents

	Page
Preface	ii
List of Figures	v
List of Tables	v
Abstract	vi
I. Introduction	1
Background: An Analysis Problem	1
Constraints on the Analysis Problem	3
Formal Problem Statement and Objective	5
Overview	6
II. Review of Dose-Response Methods	7
Dichotomous Response Models	7
One-Hit Model and Extensions	7
Probit, Logit, and Generalizations	11
Quantit Analysis	15
Aranda-Ordaz Family of Transforms	16
Time to Occurrence	18
III. The One-Hit Model	19
Determination of the Fit and Normality Assumptions	24
Conclusions and Remarks	27
IV. A Generalization of Probit and Logit	28
The Model	28
Computations for the Model	29
Computational Difficulties	33
Results for the Three Parameter Model	35
Conclusions and Remarks	38
V. Quantit Analysis	39
The Model	39
Computations for the Model	40
Results of the Quantit Analysis	42
Conclusions and Remarks	44
VI. Two Families of Transformations	46
The Symmetric Family and Associated Test	46

The Asymmetric Family and Associated Test	49
Computational Procedures	50
Results for the Symmetric Family	51
Results for the Asymmetric Family	54
Conclusions and Remarks	58
Appendix A: Description of the Avionics Evaluation Program AEP Model	60
Appendix B: Description of Mode Regression: Abort Logic . .	63
Appendix C: Sample Data	64
Appendix D: Computer Programs for Section IV	69
Appendix E: Computer Programs for Section V	74
Appendix F: Computer Programs for Section VI	90
Bibliography	101
Vita	104

List of Figures

Figure	Page
1. Probability of Abort vs MTBF.	20
2. Probability of Non-Abort vs MTBF	20

List of Tables

Table	Page
I. Chi-Square Tail Probabilities for The Probit	12
II. Chi-Square Tail Probabilities for the Logit	12
III. Chi-Square Tail Probabilities for the One-Hit Model	26
IV. Chi-Square Tail Probabilities for the Convergent Three Parameter Model	36
V. Chi-Square Tail Probabilities for the Divergent Three Parameter Model	37
VI. Chi-Square Tail Probabilities for Quantit	44
VII. Values for the Symmetric Test Statistic and MLE	52
VIII. Chi-Square Tail Probabilities for the Symmetric Family	54
IX. Values For the Asymmetric Test Statistic and MLE	55
X. Chi-Square Tail Probabilities for the Asymmetric Family.	57

Abstract

The United States Air Force has, over the past decade or so, invested much time and money in computer simulations and models. At the most basic level almost all of these simulations are input-output type procedures; variables of interest are changed to determine the effect they have on some other factor. This process is virtually indistinguishable from dose-response problems in bio-assay, hence, is capable of being analyzed by the same methods used in bio-assay. The two most commonly used techniques are probit and logit, but there are many other available techniques. An alternative to performing numerous, and sometimes redundant, simulations is to use these techniques whenever possible.

Data from the Avionics Evaluation Program (AEP) were used as the basis for estimating the probability of aircraft abort, based on the mean-time-between-failure (MTBF) of various equipment items, using four quantal assay techniques. The fits obtained from these models were compared to the more popular probit and logit results previously obtained by Dr. David Barr.

EXPLORATION OF DOSE-RESPONSE TECHNIQUES WITH SOME APPLICATIONS TO A SIMULATION PROBLEM

I INTRODUCTION

Background: An Analysis Problem

It is common for analysts to use computer modeling and simulations in problem solving. The reasons for using simulations are numerous. However, simulations usually help the analyst determine the response of some system to a change in that system's environment, or operating characteristics. In other words, simulations help the analyst describe, predict, or simply understand the behavior of a complex system under a given set of circumstances.

There is no doubt that simulations are useful, but is it always practical or necessary to use them? It is not hard to think of situations where the answer is no, and the specific situation described below is just one such example.

My last project as an analyst for the Avionics Laboratory at WPAFB was to run computer simulations of a mission analysis program (Appendix A). The project was to determine the mission effectiveness of the ATF (advanced tactical fighter) with a mixed suite of existing and conceptual (new) avionics. One of the measures of effectiveness was aircraft aborts due to failure of a particular piece of equipment, or subsystem, with varying MTBFs (mean-time-between-failure).

I performed several thousand simulations for this analysis. Unfortunately, even with such a large number of simulations all I had was a dozen or so data points (i.e. MTBF-abort pairs) for each piece of equipment. If there was a need for an abort rate at some MTBF not previously considered then either another simulation, or some estimate based on the existing data would be required.

Since it was not, and usually is not, practical to run computer models for every conceivable point it was obvious that there was a need for some curve-fitting technique. It was at this point that my working group contacted Dr. David Barr (Air Force Institute of Technology) and asked him to study the problem. He determined that there were techniques for "...estimating probabilities when given a set of relative frequencies, each obtained as the response of a system to a level of quantitative stimulus, known as probit and logit analysis" (Ref 3:1).

Probit (probability unit) analysis originated in biology and its application in that area is widely accepted. The following is the general concept of probit (parenthetical matter is my own):

An analyst is interested in the effect of some drug (failure rate) on the survival (aborts) of a large number of insects (systems). One possibility is that each insect (system) survives until a certain critical dosage (MTBF) is reached and that they all die (abort) as soon as this limit is surpassed; but that is an extreme case. It is much more plausible that the critical level varies from one insect (system) to the other according to a certain distribution. When there are many independent factors determining the critical level for each insect (system), the central limit theorem may be used to justify the choice of the normal distribution. Thus, when p is the proportion of insects (systems) killed (aborted), the analyst applies the probit transformation $y=F(p)$ and he then proceeds to express y linearly in terms of the dosage (MTBF) of the drug. (Ref34 :630).

If you were to plot these dose-response (input-output) curves with the dose on the horizontal axis and response on the vertical axis the curve is S-shaped, or sigmoid, and is similar to the cumulative distribution function.

...it seems natural to try to fit a distribution function to the points... The probability curve which came to mind first was the normal, or Gaussian, distribution function; since negative values... have no physical interpretation, it made sense to make a logarithmic transformation to the lognormal distribution. This resulted in what is known as probit analysis. A similar approach, in which the logistic growth curve is used in place of the Gaussian distribution function, is known as logit analysis (Ref 3:5).

Dr. Barr's work (Ref 3) on the problem of curve-fitting showed potential for the application of these statistical techniques to the 'input-output' type problem discussed earlier. However, while these techniques worked well for some equipment items, it worked only marginally well for others, and not at all for still others. This leads to the following question; are there other dose-response techniques available to analyze the existing data (Appendix C) which would give either better fits, or at least fit those items for which the probit and logit methods were only marginal? This question is the underlying basis for this effort.

Constraints on the Analysis Problem

Before reviewing existing dose-response techniques (Section II) it is necessary to state some constraints of the existing data which limit, or eliminate the use of certain techniques. It should be kept in mind that these constraints are limiting factors only for this effort; future

efforts may be well advised to use techniques which I could not.

The number one constraint is the amount of existing data, 152 data points total for 17 equipment items. This averages out to about nine data points for each equipment item, and for some analyses this would be sufficient. Unfortunately, 100% of this data yields abort rates less than 3%; this is known as the low-dose range (usually considered < 0.10) and presents problems of its own for extrapolation over the rest of the curve.

While it would have been convenient to run more simulations to obtain a wider range of abort rates, it is no longer possible to do so. The AEP model (Appendix A) was removed from use just prior to this effort. But even before removal the model had undergone extensive modification and enhancing, which would have made any comparison of new and old data suspect.

The second constraint, while not a problem, eliminates the use of techniques known as mutihit and multistage (Refs 2, 17:1277-1278). These models will be discussed in Section II. Basically, however, the problem is that unlike living organisms, which can be exposed to a substance then periodically re-exposed (rehit) until a tumor or other response is obtained, a hardware item, within the AEP model, fails based only on its designed (one time) MTBF.

A third constraint again stems from the fact that performing new simulations is not possible, which eliminates the use of sequential methods. As the name implies, this technique involves running a simulation, or experiment, and then performing another simulation at either a higher or lower level of the stimulus based on the previous results. You then repeat this procedure until obtaining the results or accuracy

desired (Refs 8:1-2, 34).

The last set of techniques which cannot be used are known as "time to occurrence, or time to response" (Refs 9:161-163, 17:1284-1286). Since data is not available concerning when items failed there is no ready data base on which to test these techniques.

Formal Problem Statement and Objective

In a preliminary analysis effort performed by myself (Ref 26) for the Avionics Laboratory it was of interest to determine the effect of various avionics equipment MTBFs on the abort rates for the ATF. Due to time considerations, complexity of the simulation model, and working group resources it was desired to find some mathematical technique for predicting abort rates over a range of MTBFs, using a limited number of simulations for each equipment item.

A study conducted by Dr. Barr (Ref 3) showed the applicability of using quantal (dose-response) assay techniques for determining these abort rates. In particular, he was able to fit most of the 17 equipment items under consideration using the probit and logit models.

It is the purpose of this effort to determine what other quantal assay methods exist which may fit the avionics equipment in question to dose-response curves. More specifically: what quantal models are available that will estimate the probability of an abort given an MTBF for a particular piece of avionics equipment using the existing data of Appendix C? The answers to this question have implications for analysts in general, particularly where the analysts' situation involves the use of simulations.

Overview

Section II contains a summary of the types of existing quantal (dose-response) models. A brief description of each type is given along with the mathematical development where applicable.

In Sections III thru VI the one-hit, a generalization of the probit and logit, quantal, and a symmetric and asymmetric transformation are used to analyze and fit the existing sample data. In these sections the models are explained along with techniques for implementation. Also discussed are special considerations and limitations of these models as well as results and interpretation of the results.

II Review of Dose-Response Methods

Any system that yields a response to a given stimulus can use the methods listed below. The almost exclusive use of these methods in bio-assay is no reason for their non-use in other areas. The only real constraint is that the system response variable must be a random variable that takes on only one of two values; success or failure, 0 or 1, yes or no, etc., "...such observations are called binary; an older term is quantal" (Ref 12:1).

There are numerous quantitative theories that attempt to relate the frequency of response to the level of stimulus. Crump (Ref 9) and Fishbein (Ref 17) categorize the methods most commonly used into two major types: dichotomous response models and time-to-response models.

Dichotomous Response Models

One-Hit Model and Extensions. The most elementary dose-response model is the one-hit, or linear, model. "The one-hit model is obtained by assuming that, with the exception of λd hits at dosage d , the probability of exactly x hits is given by the general term of the Poisson distribution..." (Refs 17:1277, 33). The general term is as follows

$$P(X=x) = (\lambda d)^x [\exp(-\lambda d)] / x! \quad (2.1)$$

Clearly if, as in our case, only one hit is required to produce a response then the above collapses to

$$P(d) = P(x \geq 1) = 1 - P(x < 1) = 1 - \exp(-\lambda d) \quad (2.2)$$

where λ is the process rate, or rate of change of the dose response curve at $d=0$.

The point was made by Fishbein (Ref 17) that if you are working in the low dose region ($p < 0.1$) then λd is small and $P(d) = \lambda d$. This implies a simple linear model where the response is directly proportional to the dose, with slope λ . Since all of the sample data (MTBF-abort pairs) is well within the low dose region this is one possible model to use.

The natural extension to the one-hit model is the multihit model. This model considers that if at least K hits of a dose d are required to produce a response then (Ref 17:1277)

$$P(d) = 1 - \sum_{i=0}^{K-1} \frac{(\lambda d)^i \exp(-\lambda d)}{i!}$$

Note, if K is allowed to take on non-integer values then

$$P(d) = \int_0^{\lambda d} \frac{t^{K-1} \exp(-t)}{(K-1)!} dt$$

Rewriting this as:

$$P(d) = P(d; K, \lambda) = \int_0^{\lambda d} \frac{\lambda^K t^{K-1} \exp(-\lambda t)}{\Gamma(K)} dt$$

yields the generalized multihit dose-response model, or more simply a gamma distribution with scale parameter λ and a shape parameter K (Ref 33:342).

Another extension or further generalization, if you like, of this stochastic process is the multistage model, "... where the lifetime

probability of tumor induction can be expressed as

$$P(d) = 1 - \exp[-(\alpha_1 - \beta_1 d)(\alpha_2 - \beta_2 d) \dots (\alpha_K - \beta_K d)] \quad (2.3)$$

where $\alpha_i \geq 0$, $\beta_i \geq 0$, and K represents the number of transitions or mutational stages in the carcinogenic process" (Refs 17:1278, 2). This model has no application to our hardware items since the dose (MTBF) is applied only once (as an input parameter, which remains fixed, to the AEP model, see Appendix A), and where the above would collapse to:

$$P(d) = 1 - \exp[-(\alpha - \beta d)]$$

However, the multistage model could be useful if you thought about failures of an item which did not cause an abort as the mutation. Then after repairing the failed item it is replaced in the aircraft and the aircraft is flown again (note that the repair restores the MTBF and is the 'rehit' of the item). This process is repeated until a failure (the number of failures would have to be determined by some stochastic process) causes an abort (tumor).

In an article by Guess and Crump (Ref 19) there is one multistage model worth separate discussion. It is a general polynomial model for dealing with low-dose extrapolation, and is the only multistage model which is well documented and supported in a series of articles (by Guess and Crump).

To obtain this model first consider the following:

$$P(d) = 1 - \exp\left[-\prod_{i=1}^{K_1} (\alpha_i + \beta_i (d)^{l_i})\right]$$

...it is assumed that $K_1 \geq 1$ different events must occur in a single cell before a cancer (response) is

initiated ... α_i is the rate at which an initiating event for the i th stage occurs due to a spontaneous (i.e. not dose related) carcinogenesis, and β_i is the per i th power dose at which an initiating event for the i th stage occurs due to dose-related carcinogenesis (Ref 19:17).

Now rewriting the above the model becomes

$$P_f(d) \equiv P(d) = 1 - \exp(-f(d)) \quad (2.4)$$

where

$$f(d) = \sum_{i=0}^K f_i d^i \quad f_i \geq 0$$

and where f is a polynomial with nonnegative coefficients. K is the degree of f , and along with the coefficients of f must be estimated (Ref 19:18).

Note the similarity of Eq (2.4) with the general multistage model in Eq (2.3). This model, however, uses a general polynomial of unknown degree to fit the data. This model is of particular interest if one is working in the medium, or high-dose range and then wishes to extrapolate to the low-dose range for risk estimates (Ref 19:21-22).

To construct confidence intervals Crump et. al. have

...developed 'envelope curves' which are constructed for both risk and dose ... these curves are constructed by binomially simulating 100 sets of dichotomous dose-response data, representing 100 independent replications of the same experiment ... (with the same set of test doses... (Ref 10:440).

Guess and Crump (Ref 20) also develop maximum likelihood estimation techniques for this polynomial model.

Probit, Logit, and Generalizations. The two most commonly used methods of analysis of binary data are probit (Ref 15) and logit (Ref 5, 12). For the probit (log-probit) model the probability of a response being induced by a stimulus (dose) d is given by

$$P(d) = \Phi(\alpha + \beta \log d) \quad (2.5)$$

where Φ denotes the standard cumulative Gaussian (normal) distribution.

The logit model like the probit model leads to an S-shaped dose-response curve; its equation is

$$P(d) = 1/[1 + \exp(-(\alpha + \beta \log d))] \quad (2.6)$$

It approaches zero response as dose d decreases more slowly than does the probit curves since $\lim(P(d)/d^\beta) = \text{constant}$ as $d \rightarrow 0$ (Ref 17:1279).

Dr. Barr's work (Ref 3) on the curve-fitting problem, using probit and logit, yielded the values in Tables I and II for goodness-of-fit based on chi-square tail probabilities for the 17 equipment items considered. These values will be the bench mark against which all other model fits will be tested (primarily since the probit and logit models are well developed and widely accepted).

These fits for the probit and logit are not really very different. This is somewhat expected since according to Finney the logistic and normal distributions are "... scarcely distinguishable ... between response rates of 0.01 and 0.99..." (Ref 16:406).

There is a method described by Chambers and Cox (Ref 8) which may better discriminate between the logit and probit models. However, it depends on having a few dose (MTBF) levels then performing a test which

Table I
Chi-Square Tail Probabilities
for the Probit

<u>Item</u>	<u>Probit</u>
Bus	.9564
IMFK	.9478
CDRT	.8926
MFK	.8798
SMRT	.7812
MMP	.7784
Processors	.5844
ART	.4558
SLU	.3995
MPDG	.3674
MPDS	.3592
DEK	.1614
DSMU	.1393
HUD	.1353
INS	.0834
SCU	.0445
MTU	.0097

Table II
Chi-Square Tail Probabilities
for the Logit

<u>Item</u>	<u>Logit</u>
IMFK	.9754
CDRT	.9728
Bus	.9641
MPDG	.9075
SMRT	.8253
Processors	.8241
MFK	.8023
MMP	.7119
SLU	.4527
ART	.4284
MPDS	.3686
DEK	.3617
HUD	.1436
INS	.0668
DSMU	.0649
SCU	.0738
MTU	.0487

will determine the appropriate spacing of the dose (MTBF) levels for this discrimination. Performing more simulations is impossible at this time, hence this method cannot be used.

The vast majority of the literature encountered in this review was in the biological sciences, and thus there was a propensity to find the LD50. LD50 is known as the 50% lethal dose, or the median effective dose. While this effort has no interest in the LD50 (LD01 or less would be more informative for design engineers) it is worth mentioning since it brings up the matter of transformations.

Finding the estimates of μ , and σ^2 for the distribution given by Eq(2.5) is generally by means of the probit transformation of the experimental results. "The probit of the proportion P is defined as the abscissa which corresponds to a probability P in a normal distribution with mean 5 and variance 1" (Ref 15:21). That is, the probit of P is Y, where

$$P = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{Y-5} \exp(-u^2/2) du \quad (2.7)$$

This transformation from proportions to probits has the effect of straightening out the normal S-shaped curve. Comparing Eqs (2.5) and (2.7) shows that the probit Y is related to dosage d by the simple linear equation $Y=5+(d-\mu)/\sigma$ (Ref 15), and now to estimate LD50 you simply find the value of d which gives $Y=5$. The usefulness of this transformation is to simplify mathematical calculations.

There are numerous transformation techniques available (Refs 4, 6, 22). They depend solely on the models used, the experimental data, computational considerations, convenience of their use, and ability to in-

interpret results meaningfully.

A generalization of the probit and logit models is mentioned in an earlier paper by Prentice (Ref 31) and completely detailed by Prentice in a latter paper (Ref 30). This model takes the form of

$$P(d) = \int_{-\infty}^y f(w) dw \quad (2.8)$$

where $y = (d - \mu) / \sigma$ and μ, σ are to be estimated. The pdf, $f(w)$, has the following form

$$f(w) = \frac{\exp(wm) (1 + \exp w)^{-(m+n)}}{\beta(m, n)} \quad (2.9)$$

where β is the beta function. The logistic model, Eq (2.6), is given by $m=n=1$ and converges to the normal distribution as $m, n \rightarrow \infty$. Other special cases for various m, n values are also given (Ref 30:762). This model will be discussed more fully in Section IV where it is applied to the sample data.

Note this model is really nothing more than a beta of the second kind distribution with the transformation $u = \exp(w)$, $0 < u$. To show this note that

$$\begin{aligned} f(u) &= f(w) |dw/du| \\ &= u^m / [(1+u)^{(m+n)} \beta(m, n) u] \\ &= u^{(m-1)} / [(1+u)^{(m+n)} \beta(m, n)] \end{aligned}$$

or a beta of the second kind.

Quantit Analysis. In a paper by Copenhaver and Milke (Ref 11) a new technique is offered for analyzing quantal responses, which they call Quantit analysis. The underlying distribution is called the omega distribution by Copenhaver and Milke and is characterized by a cdf of

$$F(x(q))=q \quad (2.10)$$

and a pdf of

$$f(x(q))=1-|2q-1|^{\nu+1}$$

where $0 < q < 1$ and $\nu > -1$, and where

$$x(q) = \int_{1/2}^q dz / f(x(z)) \quad (2.11)$$

As noted by the authors this distribution is a double exponential when $\nu=0$, logistic when $\nu=1$, and uniform for the limiting case as $\nu \rightarrow \infty$.

If we let P be the probability (or proportion) of response(s) at dosage x_i , and we let $P_i = F(\alpha + \beta x_i)$ then the tolerance distribution is given by

$$f(\alpha + \beta x_i) = 1 - |2p_i - 1|^{\nu+1} \quad (2.12)$$

and

$$\alpha + \beta x_i = h_{\nu}(p_i) = \int_{1/2}^{p_i} 1 / (1 - |2z - 1|^{\nu+1}) dz \quad (2.13)$$

$h_{\nu}(p_i)$ is termed the "quantit" of p_i (Ref 11:178). The computations to obtain the parameters are similar to that for the model presented by Prentice which is discussed in Section IV. Section V, however, will discuss more fully the individual computations for the Quantit model,

and the application of it to our sample data.

Aranda-Ordaz Family of Transforms. One of the reasons for first using the logit model in the analysis of binary data is its simplicity, since the logistic transformation used to obtain the logit method is a linear function of the parameters.

The logistic transform is simple, but is it adequate? Aranda-Ordaz (Ref 1) develops simple test procedures to determine both symmetric and asymmetric departures from the logistic distribution. He also develops two families of power transforms which are alternatives to the logistic. These symmetric and asymmetric alternatives each include the logistic as a special cases.

The symmetric family, which essentially yields the same results when working with either successes or failures (Ref 1:358), is given by Equation (2.14)

$$T_{\lambda}(\theta) = (2/\lambda) \left(\frac{\theta^{\lambda} - (1-\theta)^{\lambda}}{\theta^{\lambda} + (1-\theta)^{\lambda}} \right) \quad (2.14)$$

and in the limit is the logistic when $\lambda=0$ and a simple linear for $\lambda=1$.

Now solving Eq (2.14) as a function of τ yields:

$$\theta(\tau) = \begin{cases} 0 & (\lambda\tau \leq -1) \\ \frac{(1+\lambda\tau/2)^{1/\lambda}}{(1+\lambda\tau/2)^{1/\lambda} + (1-\lambda\tau/2)^{1/\lambda}} & |\lambda\tau/2| < 1 \\ 1 & (\lambda\tau \geq 1) \end{cases} \quad (2.15)$$

We assume that τ has a linear expression in terms of some parameters associated with the explanatory variables considered in a specific situation... If we

fit by maximum likelihood a linear expression for τ for a range of values of λ we may consider the maximized log likelihood as a function of λ and hence derive not only the maximum likelihood estimate $\hat{\lambda}$, but also determine which values of $\hat{\lambda}$ provide an acceptable fit. (Ref 1:358).

The aysmmetric model is given by the family (assuming $\log W(\theta)=\tau$):

$$W_{\lambda}(\theta) = ((1-\theta)^{-\lambda}-1)/\lambda \quad (2.16)$$

This is the logistic model when $\lambda=1$, and again solving as a function of τ we get

$$\theta(\tau) = \begin{cases} 1-(1+\lambda \exp(\tau))^{-1/\lambda} & (\lambda \exp(\tau)) > -1 \\ 1 & \text{otherwise} \end{cases}$$

The same assumption about τ , and the same procedure as above is used to obtain values for λ .

There are tests (Ref 1:360-361) to determine if there are any symmetric or asymmetric departures from the logistic. The attractiveness of these tests are that they can be conducted using "... (the values) computed from the output of the logistics fit." (Ref 1:360). Another by-product of these tests, specifically the symmetric test, is that it may permit "... discrimination between the logistic and probit models..." (Ref 1:361).

The above models and tests will be fully discussed and implemented in Section VI. In that section I will try to discriminate between the logistic and probit fits already performed by Dr. Barr (Ref 3).

Time to Occurrence

According to Fishbein "The second type of dose-response modeling that is receiving increasing attention deals with the distribution of the 'time to occurrence' (latent period) and its relation to dose." (Ref 17:1284). Unfortunately the few models which were discussed by Fishbein (Ref 17) and Crump (Ref 9) all had criticisms leveled at them by the authors and others. Since these models are not well accepted, and since (as in the multistage models) new interpretations of the dose-response processes are needed this type is mentioned only for completeness.

In general this type of model could be defined as "the time of death (abort) from the type of cancer (failure) of interest or as the time of the first appearance or detection of a particular tumor type" (Ref 17:1284), parenthetical matter my own.

III The One-Hit Model

The simplest curve fitting technique, if we exclude drawing a line between two points, is the simple linear regression model. The one-hit model is nothing more than a linear regression model with a logarithmic transform of the data. Regression is a well understood and straightforward technique, and as such needs no separate discussion, except for one special consideration discussed later (regression through the origin).

If we assume that it takes exactly x failures (hits) of some item to cause an abort for a particular MTBF (dose), where the failures are independent, random events, then we can use the Poisson process Eq (2.1) to describe the probability of an abort. However, as mentioned before, if it takes only one failure to cause an abort then the general Poisson term collapses to Eq (2.2), or $P(d) = 1 - \exp(-\lambda d)$ where λ is the process rate.

It is easy to verify that if P is small ($P < 0.1$) then $P(d) \approx \lambda d$. This implies that for small P a linear equation going through the origin describes the data. While our data is much less than 10%, it comes from high MTBFs (doses). This situation is exactly the opposite of the usual bio-assay problem where a low dose causes a low response rate.

Looking at Figure 1 shows that it makes no sense to force the known portion of the curve (solid line segment) through the origin. This is because this portion of the curve has a negative slope. If we flip the curve around, as shown in Figure 2, then we might be able to force the

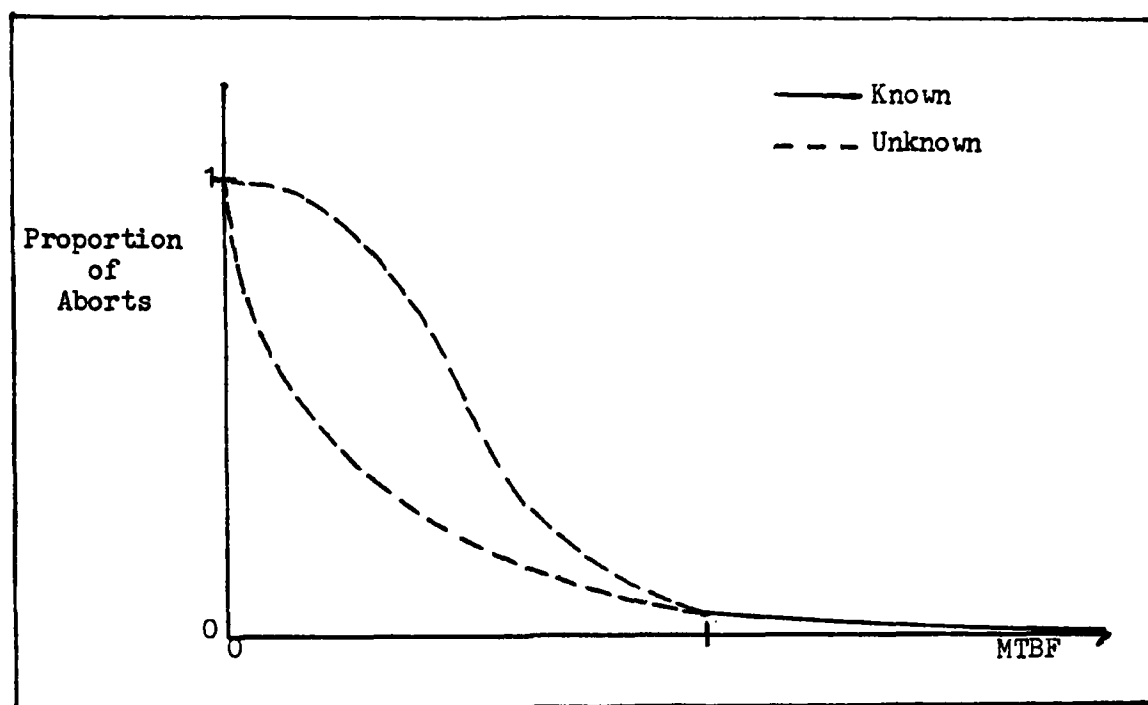


Figure 1
Probability of Abort VS MTBF

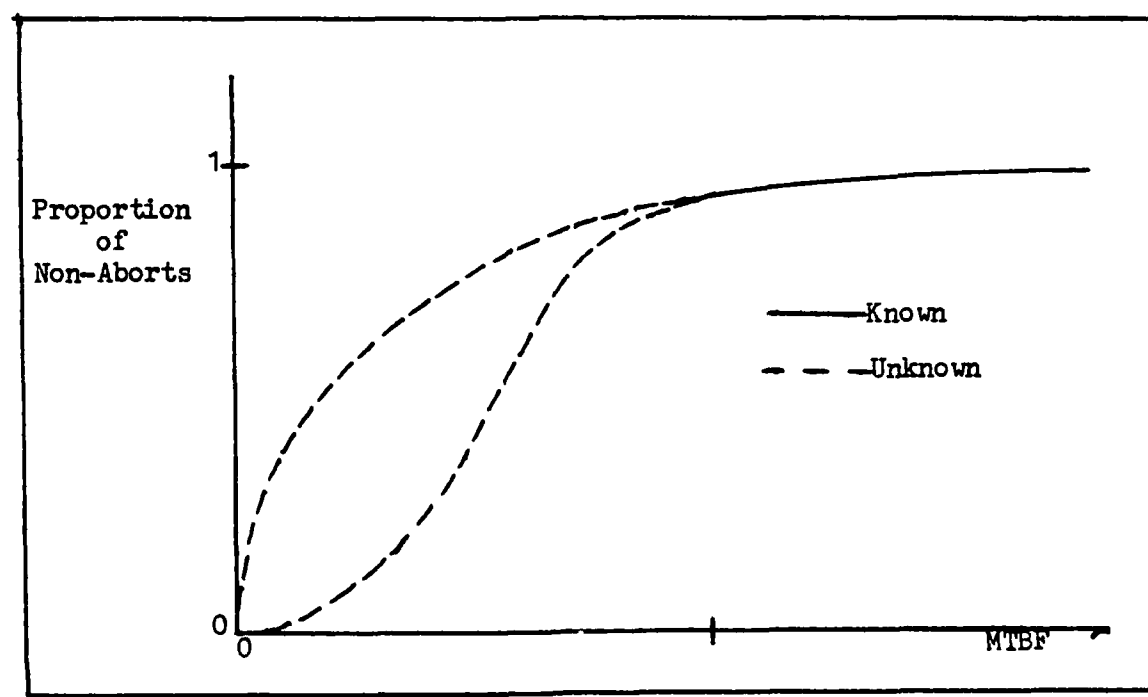


Figure 2
Probability of Non-Abort VS MTBF

Known portion of the curve through the origin and use the simple equation $P(d)=\lambda d$. However, we cannot use this method since P is no longer small and, hence, not approximated by λd . This is no handicap though, since Eq (2.2) is not hard to apply directly.

Now if we let P be the proportion of non-aborts to launches then $Q=1-P$, and Eq (2.2) can be written as:

$$Q = \exp(-\lambda d) \quad (3.1)$$

and taking logs of both sides yields $\text{Ln} Q = -\lambda d$, or

$$-\text{Ln} Q = \lambda d \quad (3.2)$$

To apply this linear model means forcing the equation through the origin, since there is no constant term. The Control Data Corporation Cyber 750 computer implementation of the Statistical Package for the Social Sciences (SPSS) (SPSS is widely available) allows one to force an equation through the origin using the appropriate option (option 19) in the regression procedure. However, each of the correlation coefficients, R^2 and adjusted R^2 , are unadjusted for the mean when using option 19. But, SPSS displays an extra line of output with these values adjusted for the mean as suggested by Theil (Ref 34:76).

When forcing an equation through the origin using SPSS one should first determine the appropriateness of this option. In our case it makes perfect sense, since we would expect to have a 'continuous- abort' for an MTBF of zero for any item which, by itself, causes an abort. Another item for consideration is the adjusted for the mean correlation

coefficients, R^2 and adjusted R^2 . If these values should happen to go negative then forcing the equation through the origin is not appropriate (Ref 34:177).

After performing an initial regression on all 17 equipment items (Appendix C), and observing that the correlation coefficients adjusted for the mean were all negative, a new approach was undertaken. Instead of using the MTBF the simple transform $d' = \ln \text{MTBF}$ was tried, or $-\ln q = \lambda d' = \lambda \ln d$. If we now solve for P we obtain:

$$P = 1 - (d)^{-\lambda} \quad (3.3)$$

With Eq (3.3) note that the value of the MTBF can approach but not equal zero. Since there is no reason not to allow the MTBF to equal zero, at least in theory if not actual practice, I applied another simple transformation, really a translation. The translation was to simply add one to the MTBF before taking logs, yielding:

$$P = 1 - \exp(-\lambda \ln(d+1))$$

or

$$P = 1 - (d+1)^{-\lambda} \quad (3.4)$$

Besides letting the MTBF take on all non-negative values Eq (3.4) has the property of being a known distribution, the Pareto distribution, translated by a value of one.

Using Eq (3.4) and performing regression on all 17 equipment items gave acceptable fits for all of the items. The determination of acceptability was by the usual methods; checking the F-test values, correla-

tion coefficients, residuals, etc. Listed below are the equations obtained from the regression procedure (remember P is the proportion of non-aborts).

ART	$P = 1 - \exp[-.94428329 \ln(d+1)]$
Bus	$P = 1 - \exp[-1.0278328 \ln(d+1)]$
CDRT	$P = 1 - \exp[-1.0855558 \ln(d+1)]$
DEK	$P = 1 - \exp[-.97513613 \ln(d+1)]$
DSMU	$P = 1 - \exp[-.95996496 \ln(d+1)]$
HUD	$P = 1 - \exp[-.91812855 \ln(d+1)]$
IMFK	$P = 1 - \exp[-1.1425878 \ln(d+1)]$
INS	$P = 1 - \exp[-.98828411 \ln(d+1)]$
MFK	$P = 1 - \exp[-1.0553776 \ln(d+1)]$
MMP	$P = 1 - \exp[-.92132193 \ln(d+1)]$
MPDG	$P = 1 - \exp[-1.2435246 \ln(d+1)]$
MPDS	$P = 1 - \exp[-1.2155868 \ln(d+1)]$
MTU	$P = 1 - \exp[-.88618825 \ln(d+1)]$
Processor	$P = 1 - \exp[-.84867598 \ln(d+1)]$
SCU	$P = 1 - \exp[-.93728168 \ln(d+1)]$
SLU	$P = 1 - \exp[-.84538133 \ln(d+1)]$
SMRT	$P = 1 - \exp[-.97621868 \ln(d+1)]$

Note the clear trend in the above equations for the coefficient to cluster about the value of $\lambda=1$. Dr Barr (Ref 3) using probit and logit had no such trend for any of his coefficients.

Determination of the Fit and Normality Assumptions

During the preliminary AEP analysis (Ref 26) there was a real concern about the random number generator used in the AEP model. For a few of the equipment items in Appendix C there appears to be significant departures from the general trend. There is a way to determine if there are departures, at least for the normality assumptions, when using regression. This is by performing residual analysis, which is well known and easy to perform using SPSS. However, for a detailed description of residuals and residual analysis, the reader should see Theil (Ref 34).

The initial SPSS residual plots for the 17 equipment items indicated that a few data points were possible 'outliers'. But, further investigation using a $t(n-2)$ distribution, since n was small in all cases, revealed no real outliers. The residual plots appeared to show no heteroscedasticity, or more simply the variance appeared to be constant. I emphasize the word appear, since the residual plots had as few as six points, and with so few points it would be misleading to state that there was absolutely no heteroscedasticity. However, in general, there were no significant departures from normality.

To determine how well this model compares with the probit and logit methods I used the contribution to the chi-square tail probabilities, as did Dr. Barr.

If the true probability of an abort for a given level of MTBF is known to be p , and if the number of launches is n , then the number of aborts has a binomial distribution with parameter (n,p) ; if n is large then the number of aborts can be approximated by a normal distribution with mean np and variance $np(1-p)$. (Ref 3:8, also see 14:229-230).

Hence, if the normal approximation is good for the abort process then the square of this will be chi-square distributed. That is, if we let x be the number of aborts then "... $(x-np)/[np(1-p)]^{1/2}$ will have a normal distribution..., and so $(x-np)^2/np(1-p)$ will have a chi-square distribution..." (Ref 3:8). Note, if we have K levels of the MTBF the chi-square distribution will have $K-1$ degrees of freedom (df). We lose one degree of freedom since we must estimate λ .

The above chi-square distribution, and hence the validity of using chi-square comparisons, hinges on whether or not the normal approximation of the abort process is good. I bring up this issue since the average p for our data is approximately .002, a very small value, and this implies the underlying binomial process is extremely skewed to the right. Also the theory is for large n , but how large should this n be? Our n was, on average, approximately 5300; is this large enough?

There is an excellent paper by Raff (Ref 32) with some easy to use graphs for determining the appropriateness of using a normal, Poisson, etc. approximation to a binomial process. Unfortunately his graphs were of little help in our case since p was so small, and n extended beyond the range of his graphs.

I had to insure that the normal approximation of the underlying binomial process was good, or the chi-square comparisons would be meaningless. However, the test for normality was simple. First, I generated 100 binomial random variables with $p=.0024$ and $n=5300$. Next I let $y=(x-np)/[np(1-p)]^{1/2}$, where x is the binomial random variate generated.

Then, using the Kolmogorov-Smirnov test in SPSS, I tested whether y was standard normal distributed. The results were that I could find no evidence that the abort process was not normal; hence, I could assume the chi-square comparisons would be valid. The results of the one-hit chi-square fit values is given in Table III.

Table III

Chi-Square Tail Probabilities
for the One-Hit Model

<u>Item</u>	<u>Probability</u>
Bus	.9584
CDRT	.9543
MFK	.6789
SLU	.5852
IMFK	.4935
MMP	.4438
DEK	.3855
MPDG	.2127
ART	.1984
SMRT	.1313
Processors	.1278
MPDS	.0458
INS	.0187
HUD	.0095
MTU	.0028
DSMU	.0007
SCU	.0002

Comparing Tables I and II with Table III it is seen that the one-hit model gives a better fit for the SLU than either probit or logit. Also the one-hit fits the CDRT and DEK better than the probit. However, with the exception of the SLU, no item in Table III had a better fit than that given by logit. Overall, fits using the one-hit model were

much worse than those obtained by either logit or probit.

Conclusions and Remarks

The one-hit method is simple and easy to use. Any analyst with only a programmable calculator (and even many cheaper models) can perform linear regression. The methods and theory of regression can be found in almost any elementary statistics text, as well as many other places. However, the simplicity of the one-hit linear model should not be the only reason for its use. I found that while the one-hit model fit the data well, in the regression sense, it did not fit the data well by the criterion of the chi-square tail probabilities. More simply, at least two other models fit the data better than the one-hit.

The one-hit model is a valid technique, since it is nothing more than linear regression, but it is not the only technique. Any analyst wishing to use the one-hit model should insure that it is the best technique for his/her situation.

IV A Generalization of Probit and Logit

Previous sections indicate that the most widely used and explored dichotomous response models are the probit and logit. In the literature reviewed scarcely an article ended without some mention of at least one of these models. However, discrimination between these two models is often difficult.

Some consideration (...Chambers and Cox [1967]...) has been given to the choice between probit and logit models with the general result that extremely large sample sizes are required to effectively discriminate between the two. Little success, however, seems to have been achieved in the development of sensitive tests of fit for probit and logit models or in the development of alternate classes of models when the usual models prove inadequate (Ref 30:761).

Hence, the almost indistinguishable results of the chi-square fits for probit and logit obtained by Dr. Barr (Ref 3) are not surprising (Refs 8, 16:406). "This raises the question as to whether tests based on more specific alternatives may be more sensitive" (Ref 30:762).

One method of a goodness of fit procedure is to embed the models into a more general parametric family of models and then test the specific models, using ordinary likelihood procedures, relative to the general one. (Refs 7, 30). This is exactly what Prentice (Refs 30, 31) does.

The Model

The probability of response (abort) for a given dose d is given as

$$P(d) = \int_{-\infty}^Y f(w) dw \quad (4.1)$$

where $y=(d-\mu)/\sigma$ and μ, σ are unknown and where the family is given by

$$f(w) = \frac{\exp(wm)(1 + \exp w)^{-m-n}}{\beta(m,n)} \quad (4.2)$$

where β is the beta function, and $m, n > 0$.

If $m=n=1$ then Eq(4.2) is $f(w)=\exp(w)/(1 + \exp w)^2$ or rewriting the response $P(d) = \exp(y)/(1 + \exp y)$, which is a logarithmic transformation of the beta of the second kind. Prentice (Ref 31), after some reparameterization, shows that Eq(4.2) converges to the normal distribution as $m, n \rightarrow \infty$.

Other limiting special cases are the extreme minimum value ($m=1, n \rightarrow \infty$) and extreme maximum value ($m \rightarrow \infty, n=1$) densities ... other limiting distributions are double exponential ($m \rightarrow 0, n \rightarrow 0$), exponential ($m \rightarrow 0, n \rightarrow \infty$), and reflected exponential ($m \rightarrow \infty, n \rightarrow 0$) (Ref 30:762).

The density function given by Eq(4.2) is symmetric along $m=n$ or $P(d) = 1-P(d)$. Equation (4.2) is also "negatively skewed for $m < n$ and positively skewed for $m > n$ " (Ref 30:762). Another characteristic of this density is that it either has narrower or fatter tails than the logistic depending on whether $m > 1$ or $m < 1$ respectively (Ref 7:1889). "Note that [Eq 4.2] allows the choice between alternative models, such as the probit and logit models, to be reduced to the choice between values in a single model" (Ref 30:762).

Computations for the Model

As was the case for the one-hit, and quantal response models in general, the fitting of the distribution to the resulting sigmoid re-

sponse curve is based on the conditional probability of the number of responses x given dose d , or:

$$P(x|d) = \binom{n}{x} P^x (1-P)^{n-x} \quad (4.3)$$

where d is the dosage, n is the number of individuals at dose d , x is the number of positive responses, and P is given by the distribution function.

The likelihood function of Eq(4.3) is

$$L = \prod_{i=1}^k \binom{n_i}{x_i} P^{x_i} (1-P)^{n_i-x_i} \quad (4.4)$$

Now suppose we have x_i responses for n_i individuals at dose d_i $i=1, \dots, k$. Also let $P(d)$ represent the probability of response at dose d_i , where P depends on $\underline{\theta} = (\theta_1, \dots, \theta_t)$, then the log-likelihood for $\underline{\theta}$ is simply

$$\ell = \sum_{i=1}^k [x_i \log P(d_i) + (n_i - x_i) \log Q(d_i)]$$

where $Q=1-P$.

The derivative $s = d\ell/d\theta$ has j th component

$$d\ell/d\theta_j = \sum_{i=1}^k [x_i/P(d_i) - (n_i - x_i)/Q(d_i)] [dP(d_i)/d\theta_j] \quad (4.5)$$

and the (j,h) element of the information matrix is

$$\sum_{i=1}^k [n_i/P(d_i)Q(d_i)] [dP(d_i)/d\theta_j] [dP(d_i)/d\theta_h] \quad (4.6)$$

The information (Fisher information) matrix is the matrix of negative

expected values of all second partials of the log-likelihood function or

$$I_n(\theta) = -E_{\theta}[l''(x|\theta)]$$

For more about the information matrix the reader is referred to DeGroot (Ref 14) section 7.8, and Theil (Ref 34) section 8.4.

As noted by Prentice calculations for l and its derivatives are simple provided $P(d)$ and $dP(d)/d\theta$ are easy to compute. If $P(d)$ is given by Eq(4.1) then a convenient method for computing $P(d)$, with underlying density given by Eq(4.2), is $I(z; m, n) / B(m, n)$ where $z = \exp(y) / (1 + \exp y)$ and I represents the incomplete beta integral (Ref 30:768). If we let $\theta = (\mu, \sigma, m, n)$ then "the derivative of $P(d)$ with respect to θ has straightforward components $dP(d)/d\mu = \sigma^{-1} f(y)$ and $dP(d)/d\sigma = \sigma^{-1} y f(y)$ " (Ref 30:763). However:

$$dP(d)/dm = \int_{-\infty}^y (d \log f(w)/dm) f(w) dw \quad (4.7)$$

and also

$$dP(d)/dn = \int_{-\infty}^y (d \log f(w)/dn) f(w) dw \quad (4.8)$$

may not have closed form solutions. Prentice suggests that "for any fixed (m, n) a straightforward Newton-Raphson procedure can be used to compute $\hat{\mu} = \hat{\mu}(m, n)$ and $\hat{\sigma} = \hat{\sigma}(m, n)$ " (Ref 30:763).

n Performing the calculations for μ and σ as functions of m, n are, at least in theory, quite simple. First consider

$$\theta' = (\theta_1, \theta_2) \quad \text{where} \quad \theta_1' = (\mu, \sigma) \quad \text{and} \quad \theta_2' = (m, n)$$

and

$$\underline{s}' = (\underline{s}_1, \underline{s}_2) \text{ where } \underline{s}_1' = (d\ell/d\mu, d\ell/d\sigma) \text{ and } \underline{s}_2' = (d\ell/dm, d\ell/dn)$$

Next partition the information matrix $I(\theta)$ so that $I_{11}(\theta)$ is the upper 2 by 2 block of $I(\theta)$. Next for fixed θ_2^0 "a Newton-Raphson procedure iteratively updates trial value θ_1^0 to $\theta_1^0 + I_{11}(\theta)^{-1} \underline{s}_1$, where $I_{11}(\theta)^{-1}$ and \underline{s}_1 are evaluated at θ_1^0, θ_2^0 until convergence to the MLE $\hat{\theta}_1$ is reached" (Ref 30:764). Note that the asymptotic covariance matrix of $\hat{\theta}_1$ is $I_{11}(\theta)$ evaluated at $(\hat{\theta}_1, \theta_2^0)$.

We now have a method for determining, by MLE's, the parameters (μ, σ) . However, as noted before, difficulty in calculating $P(d)$ with respect to either m or n hampers "the use of asymptotic likelihood methods for simultaneous inference on all four parameters (μ, σ, m, n) " (Ref 30: 766). The use of a grid of (m, n) values may overcome this difficulty. That is, you maximize the log-likelihood for numerous fixed values of (m, n) .

There are some convenient three parameter submodels of Eq(4.2) for which the estimation is more direct. The specific case of $n=1$ yields

$$P(d) = [\exp(y)/(1 + \exp y)]^m \quad (4.9)$$

the derivative of which is

$$dP(d)/dm = P(d) \log[\exp(y)/(1 + \exp y)]$$

Now simultaneous inference on all three parameters (μ, σ, m) is easily performed by a slight modification to the previously mentioned iterative procedure.

Computational Difficulties

In theory any of the above procedures, the grid search over (m,n) , the three parameter model, or even simultaneous solutions for all four parameters are possible using a Newton-Raphson procedure. However, the Newton procedure, in general, has problems of its own, "if the initial starting point is 'too far' from a minimum, the method will not converge" (Ref 27:443). The problem of initial estimates is widely known, but in our case, as I will explain, it is very evident.

In the article by Prentice (Ref 30) he applied the three parameter model given by Eq(4.9) to some classical insect mortality data. The model solutions for μ , σ , and m were given and this gave me an excellent chance to verify and validate my computerization of the model in Eq(4.9). However, I soon discovered how sensitive to initial starting values this method really is.

If I fixed any two of three parameters (μ, σ, m) to the results listed by Prentice and then varied the third, I found that I could only change this parameter by something less than 10%. Outside of this 10% range the model 'blew-up'. By this I mean that the computer (a Control Data Corporation Cyber 750) would either go into machine underflow or overflow, or when trying to invert the information matrix I found that the matrix was ill-conditioned. These were just a few of the problems. Remember this was just for one parameter; the problem intensified directly with the number of parameters varied. This presented a real problem since I had no idea as to the values of μ or σ for any given m within our sample data, let alone the value for m .

My initial reaction was to do a search for m , incrementing it a

little at a time; this would only leave the problem of estimating μ and σ . But, as I soon discovered, different values of the shape parameter m required different initial estimates for μ and σ . It did not take long for me to give up on this approach.

The next approach was the most costly with respect to time. The approach was to systematically patch the computer code for each problem encountered. I inserted numerous error checks, put in bounds to prevent computer over/underflow, I even modified the technique such that if the search procedure started to diverge, it would automatically move back half the distance (towards the last good solution). Unfortunately, all of this simply slowed down the divergence, but the results were the same.

In hindsight the solution was simple. Since the problem was one of finding the zeros of the derivatives of the log-likelihood function, I only needed a routine that solved nonlinear equations. The solutions from this routine would then be given to the Newton procedure as initial estimates. After examining many such techniques I found one in the IMSL package (ZXCGR) which worked well. It is a conjugate gradient algorithm for finding the minimum of a function.

The conjugate gradient technique is much more forgiving for 'bad' initial guesses, but it too has limits for these guesses. So one more technique was added to the chain of solutions. Note, if $y' = a + b \ln x = (\ln x - \mu)/\sigma$ then $1/b = \sigma$ and $\mu = -\sigma a$, but $y' = a + b \ln x$ is simply the linear regression model for the logit procedure. Using the regression coefficients from the logit procedure, to obtain initial estimates for μ, σ , for the gradient technique gave the best results thus far.

Results for the Three Parameter Model

The first step was to run a simple linear regression on the log-linear equation for logit. Then, using the coefficients from the regression, calculation of initial estimates for μ, σ was performed and given to the gradient algorithm along with an estimate of $m=1$. The results for $\hat{\mu}, \hat{\sigma}, \hat{m}$ out of the gradient procedure were then given to the Newton procedure. (See Appendix D for the computer listings for the gradient and Newton procedures.) The results of the analysis follow.

In six cases (ART, INS, MFK, MMP, and SMRT) the initial estimates out of the gradient search algorithm let the Newton procedure converge. However, the estimates for $\hat{\mu}, \hat{\sigma}, \hat{m}$ out of the Newton procedure were virtually unchanged from those input by the gradient procedure.

The equation for the three parameter model is given by Eq(4.9), where $y=(d-\mu)/\sigma$. The values of $\hat{\mu}, \hat{\sigma}, \hat{m}$ for the six equipment items which converged are given below, and the chi-square values are given in Table IV.

ART	$\mu=4.93887$	$\sigma=.83589$	$m=.81631$
DSMU	$\mu=6.31593$	$\sigma=.58915$	$m=.80487$
INS	$\mu=5.28186$	$\sigma=.81942$	$m=.81643$
MFK	$\mu=5.75742$	$\sigma=.68732$	$m=.80448$
MMP	$\mu=6.83882$	$\sigma=.72884$	$m=.80344$
SMRT	$\mu=3.5372$	$\sigma=.98229$	$m=.85481$

Table IV

Chi-Square Tail Probabilities for the
Convergent Three Parameter Model

<u>Item</u>	<u>Value</u>
MFK	.8364
MMP	.7425
SMRT	.7239
ART	.3871
DSMU	.2786
INS	.8656

For the other eleven cases the estimates out of the gradient procedure caused the Newton procedure to diverge. Even though the estimates for $\hat{\mu}, \hat{\sigma}, \hat{m}$ were not 'good' initial estimates for the Newton procedure they are listed below for completeness. The chi-square values are given in Table V.

Bus	$\mu=1.42526$	$\sigma=.66959$	$m=2.53973$
CDRT	$\mu=-.612943$	$\sigma=.84179$	$m=3.85879$
DEK	$\mu=-1.77332$	$\sigma=.94971$	$m=10.76879$
HUD	$\mu=1.38298$	$\sigma=.86383$	$m=1.20253$
IMFK	$\mu=-.85582$	$\sigma=.98331$	$m=1.01454$
MPDG	$\mu=-1.94282$	$\sigma=.90172$	$m=2.17709$
MPDS	$\mu=-2.25281$	$\sigma=1.02966$	$m=1.64836$
MTU	$\mu=-1.26420$	$\sigma=1.18823$	$m=4.74657$
Processors	$\mu=-1.29703$	$\sigma=1.10001$	$m=5.73043$
SCU	$\mu=-1.03275$	$\sigma=.87300$	$m=1.84555$
SLU	$\mu=-.91022$	$\sigma=1.15546$	$m=2.54579$

Table V

Chi-Square Tail Probabilities for the
Divergent Three Parameter Model

<u>Item</u>	<u>Value</u>
CDRT	.9483
IMFK	.9343
Bus	.9087
Processors	.7466
SLU	.2918
MPDG	.2747
DEK	.2279
MPDS	.2173
HUD	.0664
SCU	.0147
MTU	.0061

For the items which did converge, using the Newton procedure, there is a clear trend for the values of \hat{m} to be close to 0 and for $\hat{\mu}$ to be around 5. But, for those items which diverged, \hat{m} is larger than 1 and $\hat{\mu}$ is very small. (Remember μ is actually an estimate of the dose which in our case is the Log MTBF).

Comparing Table IV with Tables I and II it can be seen that only one item, the DSMU, had a better fit under the three parameter model than either logit or probit provided. Two other items, the MFK and MMP, had better fits than with logit. All fits were worse than the probit except the DSMU.

Looking at Table V shows that all of the items had worse fits than that provided by the logit, and most had worse fits than the probit. But, these items were the divergent ones and their fits are questionable anyway.

Conclusions and Remarks

The technique offered by Prentice is interesting to say the least. It contains within its family both the logistic and normal distributions which are so often used. Prentice offers a standard score test to examine the fit of the logistic and normal models which is more sensitive than the usual chi-square test of fit. This test is based on the asymptotic distribution of $\mathbf{g}' = (d\ell/dm, d\ell/dn)$ evaluated at (m, n) .

However sensitive the above test may be the computations for it depend on convergence of the Newton procedure so that the information matrix may be obtained. You must also reparameterize as m and n approach infinity for the normal model. But, I must admit, this is not really a drawback if you have good estimates of μ and σ .

The application of the three parameter model did not give better results than those obtained by the more familiar probit and logit methods, except for a few items noted previously. Even for these few items the computational techniques were extremely more complex and time consuming. However, this does not imply that the technique must always be so difficult.

I feel that a future effort, dedicated to computerizing the entire family, would be useful. There were many numerical techniques (Refs 23, 25, 27) which I did not investigate, due to time considerations, which may have been useful in overcoming the computational difficulties which I encountered.

The Model

Copenhaver and Mielke (Ref 11) give another family of distributions, analogous to that of Prentice (Ref 30). This family is designated as the omega distribution. The cdf of this distribution is

$$F(x(q)) = q \quad (5.1)$$

its pdf is

$$f(x(q)) = 1 - |2q - 1|^{v+1} \quad (5.2)$$

and

$$x(q) = \int_{1/2}^q [f(x(z))]^{-1} dz \quad (5.3)$$

where $0 < q < 1$ and $v > -1$. As was the case for the family given by Prentice, this family has embedded in it special cases. "In particular, this distribution is a double exponential when $v=0$, a logistic distribution when $v=1$, and a uniform in the limiting case as $v \rightarrow \infty$ " (Ref 11:177). That is, for $v=1$ the pdf is $f(x(q))=4q(1-q)$, and

$$x(q) = \int_{1/2}^q [4z(1-z)]^{-1} dz = (1/4) \log(q/1-q)$$

or the logit of q . This yields the logistics density function:

$$f(x) = 4 \exp(4x) [1 + \exp(4x)]^{-2}$$

For $v=0$ the double exponential density function is $f(x)=\exp(-2|x|)$, and for $v \rightarrow \infty$ the density is $f(x)=1$ for $-1/2 < x < 1/2$.

Note that the symmetric distributions of Eq(4.2) (e.g when $m=n$) given by Prentice are approximately the "...subset of omega distributions where $0 < \nu < 2$ " (Ref 11:178). But, as can be seen by Eq(5.2), the omega distribution includes no asymmetric distributions.

The likelihood function for quantit analysis is the same as that in Eq(4.3), but now

$$P_i = \int_{-\infty}^{\alpha + \beta x_i} f(t) dt = F(\alpha + \beta x_i)$$

Thus, the tolerance distribution is given by

$$f(\alpha + \beta x_i) = 1 - |2P_i - 1|^{\nu+1}$$

and

$$\alpha + \beta x_i = h_{\nu}(P_i) = \int_{1/2}^{P_i} 1/(1 - |2z - 1|^{\nu+1}) dz \quad (5.4)$$

where $h_{\nu}(P_i)$ is termed the "quantit" of P_i .

Computations for the Model

"The computational procedures for obtaining ML estimates of parameters α, β and ν ($\hat{\alpha}, \hat{\beta}, \hat{\nu}$) consists of an efficient search routine for determining $\hat{\nu}$ " (Ref 11:178). The calculations are identical to that for Prentice (see Section IV), however; the procedure is more tractable since the numerical calculations are simpler, due to 'nice' functions.

The procedure starts at $\nu=1$ and continues until $\hat{\nu}$ is of the desired accuracy or $\hat{\nu}$ exceeds 20. "If ν is 20 or larger, there is very little difference between the omega distribution and limiting uniform distribution..." (Ref 11:178). For the initial value of $\nu=1$ use as

initial estimates of α and β the least squares solution of the equation $h_v(P) = \alpha + \beta x$, where $h_v(P)$ is the observed quantile corresponding to the proportion of observed responses s , at dose x , to the number of trials at dose x .

The derivatives of the log-likelihood and the elements of the information matrix are calculated exactly as in equations 4.5 and 4.6 respectively. However, for the omega model with $P_i = F(\alpha + \beta x_i)$ the derivatives with respect to the parameters α, β are

$$dP/d\alpha = f(\alpha + \beta x)$$

and

$$dP/d\beta = x f(\alpha + \beta x)$$

So, if you have the j th iterative estimates of α_j and β_j and $(j-1)$ st iterative estimate of P_i then the Newton-Raphson procedure yields the j th estimate of P_i . The $(m+1)$ st iterative solution of $P_{i,j}$ is given by

$$P_{ij}^{(m+1)} = P_{ij}^{(m)} + (1 - |2P_{ij}^{(m)} - 1|^{v+1}) [\alpha_j + \beta_j x_i - h_v(P_{ij}^{(m)})]$$

and since we can write Eq(5.4) as

$$G(P_{ij}) = \alpha_j + \beta_j x_i - h_v(P_{ij}) = 0$$

and

$$G'(P_{ij}) = -(1 - |2P_{ij} - 1|^{v+1})^{-1}$$

hence

$$P_{ij}^{(m+1)} = P_{ij}^{(m)} - G(P_{ij}^{(m)}) / G'(P_{ij}^{(m)})$$

Since the functions, and thus the above equations, are "well-behaved" the 2 by 2 information matrix never had the problems that I encountered with Prentice's model.

The only computational difficulty arises from the need to calculate the quantit $h_v(P)$. Copenhaver and Mielke initially used the infinite series given by

$$h_v(P) = \frac{s}{2} \sum_{i=0}^{\infty} \frac{|2p+1|^{i(v+1)+1}}{i(v+1)+1}$$

where

$$s = \begin{cases} 1 & P > 1/2 \\ 0 & P = 1/2 \\ -1 & P < 1/2 \end{cases}$$

"Whenever $v \geq 1$ and $|2p-1| < 0.9$, the ... series converges rapidly. In fact... the maximum error will be 10^{-8} when the first 30 terms are summed" (Ref 11:186). However, when $v < 1$ or $|2p-1|$ is close to one (as in our case) you sum for an 'appropriate' number of terms and then add a remainder. The remainder is in the form of a continued fraction. "The continued fraction converges slowly when either v is near -1 or $|2p-1|$ is close to one" (Ref 28:222). Magnus et.al. (Ref 28) develop a closed form expression for the sum of an infinite series. Copenhaver and Mielke adopted this latter technique in their computerization (Appendix E contains the source code as written by Copenhaver and Mielke).

Results of the Quantit Analysis

The values obtained for α , β and v are given below for the 17

equipment items, where $\alpha + \beta x_i = h_v(P)$ as given by Eq(5.4), and $f(\alpha + \beta x)$ is the tolerance distribution.

ART	$v=19.00$	$\alpha=-.37446$	$\beta=-.02797$
Bus	$v=20$	$\alpha=-.33545$	$\beta=-.03560$
CDRT	$v=-.90$	$\alpha=9.05674$	$\beta=-5.91737$
DEK	$v=-.90$	$\alpha=8.40762$	$\beta=-5.23146$
DSMU	$v=20$	$\alpha=-.37875$	$\beta=-.02647$
HUD	$v=20$	$\alpha=-.35067$	$\beta=-.02975$
IMFK	$v=-.90$	$\alpha=1.69570$	$\beta=-5.06704$
INS	$v=20$	$\alpha=-.37296$	$\beta=-.02658$
MFK	$v=20$	$\alpha=-.40607$	$\beta=-.02475$
MMP	$v=20$	$\alpha=-.39221$	$\beta=-.02368$
MPDG	$v=20$	$\alpha=-.44213$	$\beta=-.02624$
MPDS	$v=20$	$\alpha=-.44539$	$\beta=-.02371$
MTU	$v=-.90$	$\alpha=8.45196$	$\beta=-4.19618$
Processors	$v=-.90$	$\alpha=8.44919$	$\beta=-4.48045$
SCU	$v=20$	$\alpha=-.35513$	$\beta=-.02874$
SLU	$v=-.90$	$\alpha=6.63052$	$\beta=-4.30218$
SMRT	$v=9.50$	$\alpha=-.30000$	$\beta=-.05270$

The value of \hat{v} for all of the above, with one exception the SMRT, is either $-.9$ or 20 . These results are somewhat surprising, since for those items with $\hat{v}=-.9$ the tails are very heavy. The clear implication is this; an item with a parameter value of $-.9$ for v has a very narrow band of critical MTBFs. For MTBFs below this narrow band the

item will always cause an abort and for MTBFs above it the item will never cause an abort. Considering that the proportion of aborts is very low in the sample data, and somewhat linear (flat), this is surprising indeed.

Now looking at Table VI and comparing these chi-square values with those in Tables I and II for probit and logit respectively, it is clear that the quantit method gave consistently better fits than the probit. With only two exceptions, the MPDG and the MTU, quantit also gave better fits than did logit.

Table VI
Chi-Square Tail Probabilities
for Quantit

<u>Item</u>	<u>Value</u>
IMFK	.9915
CDRT	.9846
Bus	.9778
MFK	.9162
SMRT	.8788
Processors	.8781
MMP	.8271
SLU	.5904
ART	.5738
MPDG	.4920
DEK	.4499
MPDS	.3994
HUD	.2111
DSMU	.1895
INS	.1680
SCU	.0815
MTU	.0255

Conclusions and Remarks

This method of quantal analysis appears to have been better than all of the preceding methods. It gave consistently better fits than

probit, one-hit, and the generalization of Prentice for all 17 equipment items. It also had better fits for 15 of the 17 items than did logit.

I found that quantit analysis was quite easy to implement and understand. The computer code, given to me by Mr. Copenhaver, would be 'easy' to modify to attain any accuracy of the parameters needed. It should be noted that the code contains a subroutine to perform probit analysis (logit is a by-product of the program; $v=1$ is the logit and is always output).

I would recommend that a future effort consider modifying the code so that instead of doing fixed iterations on v , it performs a simultaneous search for all three parameters α , β and v . However, this would then have the same sensitivity problems as Prentice's model. I would further recommend that Prentice's model be incorporated with this code, thus; giving a more comprehensive analysis tool.

VI Two Families of Transformations

Previously the logistic model has been a special case of all the other models. As a matter of fact, the determination of how well these other models performed has been a comparison of them against the logistic, via chi-square tail probabilities. However, the logistic model is only a tentative model and we need to look at how adequate it is. "If we can find a procedure which detects inadequacy, and also indicates the kind of desirable modification to the model, this is potentially useful" (Ref 1:357).

Mr. Aranda-Ordaz (Ref 1) gives us two families of models which help achieve the stated objective. These families each contain the logistic and alternatives as special cases. These models each have an associated transformation which model symmetric and asymmetric departures respectively from the logistic. Symmetric transformations are such that they lead to "...essentially the same answers if successes and failures are interchanged" (Ref 1:357).

The Symmetric Family and Associated Test

The symmetric family of alternatives to the logistic is given by

$$T_{\lambda}(\theta) = (2/\lambda) \left(\frac{\theta^{\lambda} - (1-\theta)^{\lambda}}{\theta^{\lambda} + (1-\theta)^{\lambda}} \right) \quad (6.1)$$

where θ is the probability of success and λ is the transformation parameter. If we let $T_{\lambda}(\theta) = A + B \ln x$ and then solve Eq(6.1) for θ as a

function of τ we then obtain

$$\theta(\tau) = \begin{cases} 0 & (\lambda\tau/2 \leq -1) \\ \frac{(1+\lambda\tau/2)^{1/\lambda}}{(1+\lambda\tau/2)^{1/\lambda} + (1-\lambda\tau/2)^{1/\lambda}} & |\lambda\tau/2| < 1 \\ 1 & (\lambda\tau/2 \geq 1) \end{cases} \quad (6.2)$$

Note that Eq(6.1) "...reduces to the logistic transformation in the limit when $\lambda=0$ and to the linear transformation when $\lambda=1$ " (Ref 1:358). Another feature of Eq(6.1) is

$$T_{\lambda}(\theta) = -T_{\lambda}(1-\theta) \quad ; \quad T_{\lambda}(\theta) = T_{-\lambda}(\theta)$$

or the treatment of successes and failures is symmetric.

As is the case for all quantal methods the underlying process is binomial and the likelihood function is the same as that in Eq(4.4).

The systematic part of the model is given by $\tau = X\beta$.

If we fit by maximum likelihood a linear expression for τ for a range of values of λ , we may consider the maximized log likelihood as a function of λ and hence derive not only the maximum likelihood estimate $\hat{\lambda}$, but also determine which values of λ provide an acceptable fit. (Ref 1:358)

Since it was presumptuous to assume the logistic as the model it is just as presumptuous to assume the symmetric will yield, via MLE, anything but the logistic model back. Hence, before proceeding with maximum likelihood estimation of λ , we need some test to determine if there are indeed symmetric departures from the logistic model. The hypothesis we

want to test is $H: \lambda=0$.

If we consider the parameter vector β as a nuisance parameter, and replace it by its maximum likelihood estimate under $\lambda=0$, then the test statistic is the efficient score $U(\lambda)=dI/d\lambda$. $U(\lambda)$ vanishes at $\lambda=0$, but after some reparameterizing, the score takes the following limiting form:

$$U(0) = \sum_{i=1}^m (r_i - n_i \hat{\theta}_i) \tau_i^3 / 12 \quad (6.3)$$

where $\hat{\theta}_i = 1/(1+\exp(-\tau_i))$, τ_i is the logit equation, r_i is the number of responses, and n_i is the number of trials.

Since Eq(6.3) is distributed asymptotically normal (Ref 14:363) the test may be carried out with the standardized form of Eq(6.3). The rejection of $H: \lambda=0$ is for large values of the test statistic. The variance of Eq(6.3) is given by

$$I_n(\theta) = I_{\lambda\lambda} - I_{\lambda\beta} I_{\beta\beta}^{-1} I_{\beta\lambda} \quad (6.4)$$

where $I_{\lambda\lambda}$, $I_{\lambda\beta}$, $I_{\beta\beta}$, and $I_{\beta\lambda}$ are the partition submatrices of I , Fisher's information matrix (see Section IV, in particular Eq(4.6) for more, or DeGroot (Ref 14) section 7.8). The individual components are

$$I_{\lambda\lambda} = \sum_{i=1}^m n_i d_i \tau_i^6 / 144$$

$$I_{\lambda\beta_s} = \sum_{i=1}^m n_i d_i x_s \tau_i^3 / 12 \quad (s=1, \dots, p)$$

$$I_{\beta_s \beta_r} = \sum_{i=1}^m n_i d_i x_s x_r \quad (r, s=1, \dots, p)$$

where $d_i = \hat{\theta}_i(1-\hat{\theta}_i)$, and p is the dimension of $\underline{\theta}$. In our case this dimension is 2 and the x_i vectors are the unit vectors. Note, "All values required to perform the test may be computed from the output of a logistic fit" (Ref1:360).

The Asymmetric Family and Associated Test

The asymmetric family of alternatives is given by

$$W(\theta) = [(1-\theta)^{-\lambda} - 1]/\lambda$$

and

$$\log W(\theta) = \tau \quad \text{or} \quad \exp(\theta) = W(\theta) \quad (6.5)$$

where τ is linear as before. Again if we solve for θ as a function of τ we obtain:

$$\theta(\tau) = \begin{cases} 1 - (1 + \lambda \exp(\tau))^{-1/\lambda} & (\lambda \exp(\tau) > -1) \\ 1 & \text{otherwise} \end{cases} \quad (6.6)$$

Equation (6.5) also contains special cases, specifically for $\lambda=1$ Eq(6.5) reduces to the logistic and for $\lambda=0$ we get the complementary log log model.

One must be careful when using this family since it does not treat successes and failures in a similar fashion. "There are situations where it is desirable to treat successes and failures asymmetrically. Yates (Ref 37) gives some examples" (Ref 1:358).

Just as for the symmetric model the underlying process is binomial and is as described previously. The procedure is to maximize the like-

likelihood, again just as for the symmetric. We also develop a test statistic analogous to that of the symmetric. The hypothesis is now, however, $H: \lambda \geq 1$. The test statistic is given by

$$U(1) = \sum_{i=1}^m \frac{(r_i - n_i \theta_i)}{\theta_i} (\theta_i + \log(1 - \theta_i)) \quad (6.7)$$

where $\theta_i = \exp(\tau_i) / [1 + \exp(\tau_i)]$.

The variance is given by Eq(6.4), but now the individual components are

$$\begin{aligned} I_{\lambda\lambda} &= \sum_{i=1}^m n_i (\theta_i + \log(1 - \theta_i))^2 / \exp(\tau_i) \\ I_{\lambda\beta_s} &= \sum_{i=1}^m (\theta_i + \log(1 - \theta_i)) n_i x_s (1 - \theta_i) \quad (s=1, \dots, p) \\ I_{\beta_s \beta_r} &= \sum_{i=1}^m n_i x_s x_r \theta_i (1 - \theta_i) \quad (r, s=1, \dots, p) \end{aligned}$$

again p is the dimension of β . This test also requires only values from the logistic fit. The rejection of the hypothesis is for large negative values of the standardized statistic.

Computational Procedures

The calculations for both the test statistic $U(\lambda)$ and the two families are straightforward. The calculations for $U(\lambda)$ were explicitly given in the previous paragraphs and no further explanation will be given here, except for one small discussion.

It is common practice to use weighted least squares procedures when performing logit. While Mr. Arand-Ordaz does not explicitly state that he uses weighted least squares I will assume that he did. The reason

for weighting in a least squares procedure is that the assumption of homoscedasticity (constant variance) may not be justified (Cox (Ref 12) section 2.2 and Theil (Ref 34) section 6.2). The assumption of homoscedasticity is one of the key assumptions underlying all regression analysis, and must not be ignored. However, for all quantal analysis, since the underlying process is binomial, weights are easily obtainable. Since the variance of a binomial distribution is npq , it makes sense to use this as the weighting coefficient, which is what I have done. This is also what Dr. Barr (Ref 3) has done and what is suggested by Finney (Ref 15).

The model calculations, as I said before, are straightforward once you have determined if there are symmetric or asymmetric departures from the logistic. First for a fixed λ calculate the value for Eq (6.1) or Eq (6.5), for each level of the stimulus (MTBF). Next using these values perform a least squares regression to obtain β_0 and β_1 the regression coefficients. Then, using this regression equation, calculate $\theta(\tau)$ in Eq(6.2) or Eq(6.6). Finally, using the values of $\theta(\tau)$ just calculated, determine the value of the log likelihood function. Since you are trying to maximize the likelihood function the value of λ that yields the largest value of the likelihood is the MLE estimate λ .

Note, one would normally only apply one of the two families if the test statistics indicated a rejection of the null hypothesis that the distribution is logistic. However, for comparison purposes I calculated the MLE for all 17 equipment items for both families.

Results for the Symmetric Family

Given in Table VII are the results of the normalized test statistic

computed from Eqs (6.3) and (6.4), and listed in ascending order. Also listed are the MLE values for λ .

Table VII
Values for the Symmetric
Test Statistic and MLE

<u>Item</u>	<u>Statistic</u>	<u>λ</u>
IMFK	.5805	0.
CDRT	2.0290	0.
DEK	2.6989	0.
SMRT	2.8950	0.13
Processors	2.9858	0.
MFK	4.1348	0.32
Bus	5.0072	0.
ART	5.1847	0.18
SLU	6.0953	0.
INS	6.1689	0.22
MMP	7.2325	0.33
DSMU	10.3723	0.32
MPDS	13.4003	0.24
HUD	14.4408	0.62
MPDG	15.6428	0.33
SCU	29.6729	0.88
MTU	35.8081	0.

Comparing the order of Table VII and Table II (logit) it is interesting to note that the order is somewhat the same, particularly for the tops and bottoms of the tables. This is what one would expect since, as the logistic fit becomes worse, the test statistic should indicate departures from the logistic.

Remember the null hypothesis is $H: \lambda=0$, or the distribution is not different from the logistic. Therefore, the test indicates whether or not there are symmetric departures from the logistic. The values of λ , in Table VII, are fairly consistent with the expected results as indicated by the test statistic.

I am sure you have already noticed that the values of the test sta-

tistic are larger than one would expect for a standard normal distribution. This could be due to the underlying distribution not being logistic or even shaped liked a one, but something completely unrelated to the logistic. Indeed, the underlying distribution may not be represented by the family in Eq(6.1) at all. If this is the case the test for departures is not valid. This may also explain the zero values for λ for those items with very large test values.

Listed below are the equations for τ , attained at the respective λ value, for all 17 items.

IMFK	$\tau = -.88418 - 1.00979 \ln x$
CDRT	$\tau = .49617 - 1.16095 \ln x$
DEK	$\tau = .34701 - 1.01782 \ln x$
SMRT	$\tau = .06791 - .91011 \ln x$
Processors	$\tau = .50903 - .89628 \ln x$
MFK	$\tau = -1.98938 - .44583 \ln x$
Bus	$\tau = .13606 - 1.00295 \ln x$
ART	$\tau = -.39222 - .78307 \ln x$
SLU	$\tau = .00023 - .84037 \ln x$
INS	$\tau = -.57526 - .69058 \ln x$
MMP	$\tau = -2.02101 - .37681 \ln x$
DSMU	$\tau = -1.48325 - .47681 \ln x$
MPDS	$\tau = -3.33997 - .40853 \ln x$
HUD	$\tau = -2.18344 - .13002 \ln x$
MPDG	$\tau = -2.02101 - .37681 \ln x$
SCU	$\tau = -2.12457 - .01772 \ln x$
MTU	$\tau = -.74551 - .66784 \ln x$

Listed in Table VIII are the chi-square values for the fit of the 17 equipment items using the symmetric family. The values are listed in descending order. Comparing the values of Table VIII with those of Table II the order is essentially the same. However, there is an increase in the value of the tail probabilities in Table VIII for those items which did not have an MLE $\lambda=0$ (remember if $\lambda=0$ these items are fit by the logistic, hence, the same as Table II). There are three exceptions however. These exceptions are the SMRT, ART, and INS, but their tail values (Table VIII) are almost identical with those for logit in Table II.

Table VIII
Chi-Square Tail Probabilities
for the Symmetric Family

<u>Item</u>	<u>Value</u>
IMFK	.9754
CDRT	.9728
Bus	.9641
MFK	.9389
MPDG	.9075
MMP	.8261
SMRT	.8247
Processors	.8241
MPDS	.6873
ART	.4592
SLU	.4527
DEK	.3617
SCU	.3107
HUD	.2148
DSMU	.1866
INS	.0799
MTU	.0487

Results for the Asymmetric Family

Table IX lists the results of the normalized test statistic com-

puted from Eqs(6.7) and (6.4), and are in ascending order. Also listed are the values of the MLE λ .

Table IX
Values for the Asymmetric
Test Statistic and MLE

<u>Item</u>	<u>Statistic</u>	<u>λ</u>
DEK	-1.2765	-138
CDRT	- .7589	-148
Processors	- .5834	-8
IMFK	- .2972	-267
Bus	- .0467	-781
SMRT	.1812	28
SLU	.1978	-147
MFK	.8151	227
MPDG	.9758	-1789
HUD	1.1261	518
MTU	1.3683	-531
MMP	1.4385	325
ART	1.5993	61
INS	3.1261	59
SCU	3.4855	3498
DSMU	3.7158	223
MPDS	4.8448	859

The theoretical null hypothesis in this case is $H: \lambda=1$, or the logistic. Therefore, for large values of the test statistic you will reject the null hypothesis. However, Mr. Aranda-Ordaz indicates that you will reject only for large negative values. This implies that the working hypothesis is really $H: \lambda \geq 1$. The reason for this is clear; since we are interested in asymmetric departures from the logistic we are really interested in values of λ significantly less than one. While values larger than one are also acceptable the distribution starts to lose its asymmetry and becomes more symmetric looking.

If, as I said before, I were to go by the test statistic value, as an indication whether or not to proceed with the asymmetric model, I

would not proceed. I would be hard pressed to reject the null hypothesis $H:\lambda \geq 1$ for any of the items, except possibly the DEK. But, since we are interested in how well the over all asymmetric model and test perform for our data I continued, and obtained the following equations for τ . Again these equations are attained at the MLE λ given in Table IX.

DEK	$\tau = -1.74791 - .72624 \ln x$
CDRT	$\tau = -.99838 - .94095 \ln x$
Processors	$\tau = .22093 - .85761 \ln x$
IMFK	$\tau = -2.48298 - .77771 \ln x$
Bus	$\tau = -2.24865 - .75122 \ln x$
SMRT	$\tau = .93780 - 1.09026 \ln x$
SLU	$\tau = -1.71897 - .62746 \ln x$
MFK	$\tau = 3.01833 - 1.45962 \ln x$
MPDG	$\tau = -5.44571 - .45658 \ln x$
HUD	$\tau = 12.85090 - 2.66024 \ln x$
MTU	$\tau = -3.93067 - .34990 \ln x$
MMP	$\tau = 3.97039 - 1.39584 \ln x$
ART	$\tau = 1.66151 - 1.16937 \ln x$
INS	$\tau = 2.02577 - 1.17670 \ln x$
SCU	$\tau = 56.15579 - 7.87826 \ln x$
DSMU	$\tau = 5.40580 - 1.67762 \ln x$
MPDS	$\tau = .36738 - 1.17886 \ln x$

Table X contains the chi-square tail probabilities for the fit of

the data using the asymmetric family. As with all previous chi-square tables these also are listed in descending order. Note, the order is almost identical to that of Table VIII and also Table II. With four exceptions these tail probabilities are better than those in Table VIII, hence, better than those of Table II. The four exceptions are the MFK, SMRT, SCU, and HUD.

Comparing the tail values for these four items with those in Table VIII shows that while the symmetric was better it was not significantly better. The only item of the four which was not relatively the same as the logit fit, in Table II, is the SCU. For the SCU both the symmetric and asymmetric models gave chi-square values more than four times larger than that of logit, a marked improvement.

Table X
Chi-Square Tail Probabilities
for the Asymmetric Family

<u>Item</u>	<u>Value</u>
CDRT	.9934
IMFK	.9934
BUS	.9779
MFK	.9208
MPDG	.9208
MMP	.8427
Processors	.8414
SMRT	.8225
MPDS	.7623
DEK	.6693
ART	.5000
SLU	.4928
DSMU	.3857
SCU	.3068
HUD	.1556
INS	.1182
MTU	.0828

One last interesting note concerns the test. If I had performed

the asymmetric test, with the intention of only applying the model if indicated by the test, the only item I would have tried to fit would have been the DEK. And, for this single item, I would have found an almost two-fold increase in the chi-square tail probability. But, I would have missed improved fits for many other items.

Conclusions and Remarks

I feel the numerical calculations which I performed are correct and accurate, for both the test statistic values and the models themselves. However I would be less than honest if I did not report the following discrepancy.

Mr. Arand-Ordaz applied the asymmetric test to one set of sample data. This data has a point where the responses are the entire sample (e.g. the number of insects killed equaled the number exposed to the chemical stimulus) (Ref 1:362 table 4). How he treats this point was not indicated. This point is a problem, as a matter of fact any point where all or none of the subjects respond causes problems. There are many schools of thought about what to do with a point like this, but the standard seems to be to replace $p=s/n$ by $p=(s+.5)/(n+1)$ (Ref 11:178). However, if one does not use this method than something very similar to it is usually suggested.

With the above in mind I tried several different methods, but could never attain his fitted values as stated (Ref 1:362 table 4) for the logistic model. I could get very close using the SPSS regression procedure to perform the logistic fit, but never close enough to satisfy myself. Normally I would chalk this up to not knowing how Mr. Arand-Ordaz treats the problem point. However, to calculate the test statistic you

need the logistic fit. And, I could not duplicate this either, at least without "cheating" as I will explain.

Mr. Aranda-Ordaz (Ref 1:362) reports a test value of -2.76 for the asymmetric statistic. Using the closest logistic fit I could attain (compared to his) I could not get this value (remember this value is obtained by the standardized form of $U(\lambda)$, which has a mean of zero and variance as given in equation 6.4. However, I could get his value if I divided $U(\lambda)$ not by the standard deviation, but by the variance. Since the equations (6.7 and 6.4) are easy to calculate and verify, and since the SPSS regression technique is valid, I am convinced that the value of -2.76 is wrong as reported by Mr. Arand-Ordaz.

These two sets of models were relatively easy to apply and they incorporate some of the same calculations required in Sections III and IV. Hence, these models could easily be incorporated into a larger computer package which included the families given by Prentice (Ref:31) and Copenhagen (Ref 11). The three families together include the logistic (logit), normal (probit), and numerous symmetric and asymmetric alternatives. A comprehensive quantal assay computer package has much potential for use in many areas other than bio-assay

Appendix A

Description of the Avionics Evaluation Program AEP Model

The Air-to-Ground Mission Analysis (MAP) submodel of The AEP evaluates the performance of a flight of up to four aircraft on a mission which may involve multiple targets, multiple search passes, and multiple attack passes. The aircraft proceed along an externally generated nominal trajectory through the mission phases of takeoff, navigation to the search area, search, attack, and return to base. Monte Carlo techniques are applied to mean-time-between-failure (MTBF) data for the defined avionics throughout the mission to determine which subsystem modes are functioning, restoring to back-up modes, and mission aborts as required. Target location uncertainties and navigation system performance parameters are combined to define the actual flight path relative to the true target location. The sensor ground swath for the defined search pattern is then compared to the true target location to determine if the target passes through the sensor ground area coverage. Probabilities of detection, target kill, and aircraft survival are sampled to determine which mission phases are successfully completed. The model utilizes the best mode still available for each function at the time it is to be performed.

The MAP includes a detailed model of the ground turnaround process. Preflight, thruflight, postflight maintenance, ordinance arming/loading, refueling, scheduling, and launch are modeled in terms of time require-

ments and event uncertainties.

Prior to launch of the first sortie of the day, the model sequenced each aircraft through preflight maintenance and ordinance loading. During the maintenance interval equipment items are repaired and "time-to-repair" is recorded. Prior to launch of subsequent sorties, the model sequences each aircraft through ordinance de-arming, thru-flight maintenance, refueling and ordinance loading. Repair and loading times are recorded. Upon completion of the last sortie of the day, the model sequences each aircraft through ordinance de-arming, post-flight maintenance, and refueling for the next day.

The launch subfunction represents the interval between engine start to takeoff. At this time, an additional equipment check is made to determine additional failures. To determine these launches the sortie scheduling algorithms utilizes user supplied data to manage the starting time for individual sorties on sequential days.

The preflight, thruflight, and postflight maintenance times are based on mean duration time input data. Ordinance loading, arming and de-arming times are determined in a similar manner. Refueling times are based upon an input refueling rate (lbs/min) and aircraft fuel storage capacity. The model calculates fuel utilization for each sortie to determine additional fuel requirements.

In addition to ground preparation and ground maintenance, the user must also define the in-flight mission functions along with their various parameters (e.g. nav accuracy, drift rates, ect.) and their associated suites of hardware and the hardware reliability and maintainability parameters. The in-flight functions are navigation, navigation update, communication, target acquisition, weapon delivery, general flight, tar-

get, and survivability. Note that each of these functions have numerous subfunctions (e.g. navigation has radio-aided nav, and self-contained nav subfunctions).

Once airborne, an aircraft must have one of two navigation functions working and the communications function working or the aircraft will abort. The other functions which will cause an abort are the target acquisition, weapon delivery, and some of the general in-flight functions. This latter function is one that can be used to determine additional abort conditions.

Appendix B

Description of Mode Regression:

Abort Logic

For the single aircraft mission simulation the concept of mode regression and the abort process is straight forward. For each mode within each function, two things must be defined; operating or performance characteristics, and a suite of hardware items (which also has a set of parameters reliability and maintainability) needed to perform that operation (mode) within the function.

Consider one of the functions that will cause an abort if all modes fail. This function has say 10 modes and within each mode a suite of hardware items is defined. If one of the hardware items fails in mode 1 then the aircraft regresses (moves) to mode 2; you can think of this as either a backup mode (e.g. redundant aircraft equipment items may have been defined in this mode), or a degraded mode with degraded performance characteristics. Now if a crucial hardware item fails, that is one that is needed by all modes, or if enough hardware items fail such that you have regressed through all modes, then an abort will occur. Mode regression starts at the best possible mode and regresses to the less desirable modes.

Appendix C

Sample Data

The following is a list of the 17 equipment items considered for analysis, and the raw data generated by the AEP.

ART - Avionics Remote Terminal

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
3000	5370	2
2000	5354	6
1500	5378	3
1000	5437	10
750	5252	12
694	5340	9
500	5329	17
450	5402	13
350	5400	24
250	5316	44
100	5434	78

BUS - The aircraft data bus

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
20000	5443	0
3000	5380	0
1500	5309	3
1000	5386	3
900	5417	6
800	5267	6
700	5277	6
600	5275	7
500	5236	10

CDRT - Control and Displays Remote Terminal

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
4000	5413	1
3000	5396	1
2000	5374	2
1500	5390	2
1000	5299	1
750	5385	3
694	5443	5
500	5372	6
350	5365	11
275	5351	13
225	5403	13
150	5290	23
100	5388	46

DEK - Data Entry Keyboard

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5228	1
3000	5252	2
2000	5443	7
1000	5397	6
750	5238	6
500	5295	10
250	5410	24
100	5335	73

DSMU - Display Switch Memory Unit

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5359	2
3001	5406	1
2000	5394	3
1500	5300	3
1120	5443	8
900	5240	6
749	5271	12
560	5373	12
450	5308	30
350	5369	34
100	5375	74

HUD - Heads Up Display

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
2000	5313	2
1010	5334	10
1000	5326	18
750	5261	13
500	5295	21
300	5443	44

IMFK - Integrated Multi-Function Keyboard

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
3000	5333	1
2000	5472	1
750	5260	2
560	5329	4
350	5301	6
100	5204	18
75	5315	30

INS - Inertial Navigation System

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
2000	5349	8
1500	5446	5
1000	5358	13
750	5519	8
500	5364	21
250	5427	47
100	5484	115
74	5443	110

MFK - Multi-Function Keyboard

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
3000	5333	1
2000	5472	1
750	5260	8
560	5329	7
350	5301	15
100	5204	41
75	5315	54

MMP - Master Mode Panel

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
8000	5301	1
4000	5286	2
3000	5358	5
2000	5277	2
1500	5372	8
1000	5328	14
750	5344	15
500	5288	22
250	5381	36

MPDG - Multi-Purpose Display Generator

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5309	0
3000	5381	0
2000	5361	0
1500	5429	0
1120	5443	2
900	5268	0
750	5298	2
560	5274	0
450	5341	3
350	5299	1
100	5311	8

MPDS - Multi-Purpose Displays

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5356	0
3000	5270	0
2000	5412	1
1500	5443	0
1000	5386	0
750	5305	1
500	5416	2
350	5264	6
250	5333	8
100	5329	14
75	5226	9

MTU - Multiplex Terminal Unit

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
17681	5443	7
16000	5389	0
8000	5273	3
4000	5383	7
2000	5264	10
1500	5237	22
1000	5303	27

Processors - Aircraft data processors

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5240	1
4500	5363	6
4000	5272	6
3000	5443	10
2000	5365	10
1500	5379	11
1250	5433	16
1000	5309	18
900	5396	16
750	5283	18
500	5387	35
350	5307	47
100	5322	141

SCU - Sensor Control Unit

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
3973	5443	0
1500	5294	10
1000	5361	18
750	5298	11
500	5389	25

SLU - Stores Logic Unit

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
3000	5426	9
2000	5535	7
1000	5324	12
853	5443	16
694	5395	26
347	5271	39

SMRT - Stores Management Remote Terminal

<u>MTBF</u>	<u>Launches Attempted</u>	<u>Total Aborts</u>
5000	5514	1
3000	5401	1
2000	5541	3
1500	5514	8
1000	5383	8
694	5556	9
500	5498	12
347	5443	22
250	5494	34
100	5461	79

APPENDIX D

Computer Programs for Section IV

The first program, ESTM5, calculates estimates for the parameters μ , σ , and m to be output for later use by program NEWTON. This program uses a conjugate gradient search algorithm to determine the zeros of the derivative vectors of the log likelihood equations given in Section IV. The actual algorithm is the IMSL subroutine ZXCGR.

The input should be on TAPE 9 in the following form: dose, number of subjects, and number of positive responses. The very first line image on TAPE 9 should only contain the number of dose levels. The user is then asked to input initial estimates for μ , σ , and m by the interactive system. The output is written on TAPE 7 for direct use by the program NEWTON, or preliminary inspection.

The second code listing is for the program NEWTON. This program uses a Newton-Raphson method for finding the zeros of the derivative vectors of the log likelihood functions. The program expects the input data to be on TAPE 7. The input data should be in the following order:

- 1.) First card - N , the number of dose levels
- 2.) Second through $N+1$ st card - dose, number of subjects, number of responses
- 3.) $N+2$ nd card - initial estimates of μ , σ , and m .

The final estimates of μ , σ , and m along with the value of the log likelihood function at maximum is written to TAPE 8. Also on TAPE 8 is the information matrix evaluated at the final estimates.

```

1=  PROGRAM ESTMS *** THE CONJUGATE GRADIENT SEARCH METHOD***
2=  EXTERNAL FUNLH
3=  REAL X(3),G(3),FLH,W(18)
4=  COMMON/DATA/A(20,3),INUM,Z(20),DIFF(20)
5=  READ(9,'(I3)')INUM
6=  DO 222 IN=1,INUM
7=      READ(9,*)A(IN,1),A(IN,2),A(IN,3)
8=      Z(IN)=ALOG(A(IN,1))
9=      DIFF(IN)=A(IN,2)-A(IN,3)
10=222 CONTINUE
11=  ACC=.00000001
12=  DO 111 I=1,3
13=      PRINT*, '?'
14=      READ*,X(I)
15=111 CONTINUE
16=  CALL ZICGR(FUNLH,3,ACC,500,1,X,G,FLH,W,IER)
17=  CALL FUNLH(IDUM,X,FDUM,GDUM,-999)
18=  PRINT*,FLH,X(1),X(2),X(3)
19=  PRINT*,G(1),G(2),G(3)
20=  END
21=  SUBROUTINE FUNLH(X,F,S,FLAG)
22=  REAL X(3),S(3),M2,F,P(20),DMU(20),DSIGMA(20),DM1(20)
23=  COMMON/DATA/A(20,3),INUM,Z(20),DIFF(20)
24=  M2=1.0
25=  IF(FLAG.NE.-999)THEN
26=5   G1=GAMMA(X(3))
27=   G2=GAMMA(M2)
28=   G3=GAMMA(X(3)+M2)
29=   BETA=G1*G2/G3
30=*****
31=***** THIS DO LOOP CALCULATES THE CDF (P(J)) FOR EACH DATA POINT
32=***** AND THE DERIVATIVE WRT EACH OF THE PARAMETERS, MU,SIGMA, M 1,
33=***** THE CDF FUNCTION ***
34=*****
35=  DO 10 J=1,INUM
36=      Y=(Z(J)-X(1))/X(2)
37=      TEMP=EXP(Y)/(1+EXP(Y))
38=      IF(X(3).LT. 0.0) THEN
39=          P(J)=.9999999999
40=      ELSE
41=          P(J)=TEMP**X(3)
42=      END IF
43=      DM1(J)=P(J)*ALOG(TEMP)
44=      PDF=(EXP(Y*X(3))*(1+EXP(Y))**(-X(3)-M2))/BETA
45=      DMU(J)=PDF/(-X(2))
46=      DSIGMA(J)=Y*PDF/(-X(2))
47=10 CONTINUE
48=  DO 20 I=1,3
49=      S(I)=0.0
50=20 CONTINUE

```



```

51=*****
52=*** THIS LOOP CALCULATES THE DERIVATIVE VECTOR S(I) OF THE LOG-
    LIKELY
53=*****
54= DO 40 K=1,INUM
55=     TEMP1=(DIFF(K)-A(K,2)*P(K))/(P(K)*(1-P(K)))
56=     S(1)=S(1)+TEMP1*DMU(K)
57=     S(2)=S(2)+TEMP1*DSIGMA(K)
58=     S(3)=S(3)+TEMP1*DM1(K)
59=40 CONTINUE
60=     FUNLL=0.0
61=     DO 60 I=1,INUM
62=     FUNLL=FUNLL+DIFF(I)*ALOG(P(I)/(1-P(I)))+A(I,2)*ALOG(1-P(I))
63=60 CONTINUE
64=     F=-FUNLL
65=     S(1)=-S(1)
66=     S(2)=-S(2)
67=     S(3)=-S(3)
68=     RETURN
69= ELSE
70=     WRITE(7, '(I3)' )INUM
71=     DO 100 KK=1,INUM
72=         WRITE(7,*)Z(KK),A(KK,2),A(KK,3)
73=100 CONTINUE
74=         WRITE(7,*)X(1),X(2),X(3)
75=     END IF
76=     RETURN
77= END

```

```

1=  PROGRAM NEWTON
2=  REAL A(20,3),P(20),DMU(20),DSIGMA(20),DM1(20),S(3),XINFO(3,3),
3=  +XINV(3,3),WK(6),X(3),M2,SS(3)
4=  READ(7,'(I3)'NDATA
5=  DO 100 IN=1,NDATA
6=    READ(7,*)A(IN,1),A(IN,2),A(IN,3)
7=100 CONTINUE
8=  READ(7,*)X(1),X(2),X(3)
9=  EPSLON=.000001
10=  M2=1.0
11=5  G1=GAMMA(X(3))
12=  G2=GAMMA(M2)
13=  G3=GAMMA(X(3)+M2)
14=  BETA=G1*G2/G3
15=  DO 10 J=1,NDATA
16=    Y=(A(J,1)-X(1))/X(2)
17=    TEMP=EXP(Y)/(1+EXP(Y))
18=    IF(X(3).LT.0.0) THEN
19=      P(J)=.9999999999
20=    ELSE
21=      P(J)=TEMP**X(3)
22=    END IF
23=    DM1(J)=P(J)*ALOG(TEMP)
24=    PDF=(EXP(Y*X(3))*(1+EXP(Y))**(-X(3)-M2))/BETA
25=    DMU(J)=PDF/(-X(2))
26=    DSIGMA(J)=Y*PDF/(-X(2))
27=10 CONTINUE
28=  DO 20 I=1,3
29=    S(I)=0.0
30=    DO 30 J=1,3
31=      XINFO(I,J)=0.0
32=30 CONTINUE
33=20 CONTINUE
34=  DO 40 K=1,NDATA
35=    TEMP1=((A(K,2)-A(K,3))-A(K,2)*P(K))/(P(K)*(1-P(K)))
36=    S(1)=S(1)+TEMP1*DMU(K)
37=    S(2)=S(2)+TEMP1*DSIGMA(K)
38=    S(3)=S(3)+TEMP1*DM1(K)
39=    TEMP2=A(K,2)/(P(K)*(1-P(K)))
40=    XINFO(1,1)=XINFO(1,1)+TEMP2*(DMU(K)**2)
41=    XINFO(1,2)=XINFO(1,2)+TEMP2*((DMU(K)*DSIGMA(K)))
42=    XINFO(1,3)=XINFO(1,3)+TEMP2*((DMU(K)*DM1(K)))
43=    XINFO(2,2)=XINFO(2,2)+TEMP2*(DSIGMA(K)**2)
44=    XINFO(2,3)=XINFO(2,3)+TEMP2*((DSIGMA(K)*DM1(K)))
45=    XINFO(3,3)=XINFO(3,3)+TEMP2*(DM1(K)**2)
46=40 CONTINUE
47=  XINFO(2,1)=XINFO(1,2)
48=  XINFO(3,1)=XINFO(1,3)
49=  XINFO(3,2)=XINFO(2,3)
50=  TEST=SQRT(S(1)**2+S(2)**2+S(3)**2)
51=  IF(TEST.GT.EPSLON) THEN
52=    CALL LGINFXINFO,3,3,3,0.0,XINV,3,SS,WK,IER)
53=    T1=0.0

```

```

54=      T2=0.0
55=      T3=0.0
56=      DO 50 I=1,3
57=          T1=T1+XINV(1,I)*S(I)
58=          T2=T2+XINV(2,I)*S(I)
59=          T3=T3+XINV(3,I)*S(I)
60=50    CONTINUE
61=      X(1)=X(1)+T1
62=      X(2)=X(2)+T2
63=      X(3)=X(3)+T3
64=      GOTO 5
65=      ELSE
66=          FUNLL=0.0
67=          DO 60 I=1,NDATA
68=              ZZ=A(I,2)-A(I,3)
69=              FUNLL=FUNLL+ZZ*ALOG(P(I)/(1-P(I)))+A(I,2)*ALOG(1-P(I))
70=60    CONTINUE
71=      WRITE(8,*)'THE LOG-LIKE FUN AT MAX= ',FUNLL
72=      WRITE(8,*)'EST OF MU= ',X(1)
73=      WRITE(8,*)'EST OF SIGMA= ',X(2)
74=      WRITE(8,*)'EST OF M1= ',X(3)
75=      WRITE(8,*)(' THE FOLLOWING IS THE INFORMATION MATRIX"')
76=      DO 80 I=1,3
77=          WRITE(8,*)XINV(I,1),XINV(I,2),XINV(I,3)
78=80    CONTINUE
79=      END IF
80=      END

```

APPENDIX E

Computer Programs for Section V

This program is for performing the calculations of the Quantit model of Section V. This code is almost exactly as given to me by Mr. Copenhagen (Ref 11), only a few changes have been made. The changes made were those needed in order to implement it on the Cyber 750 computer at WPAFB, and one small change (lines 776-781) to handle data larger than anticipated by Mr. Copenhagen. Copies of the user's manual, also obtained from Mr. Copenhagen, are obtainable from Dr. Barr in the Department of Mathematics, AFIT.

```

1=  PROGRAM QUNTIT
2=C**  QUANTIT ANALYSIS
3=C
4=C***THIS IS THE SECOND FORTRAN VERSION OF QUANTIT ANALYSIS, WRITTEN FOR
5=C  THE IBM 370/159. THIS PROGRAM WAS CONVERTED FROM THE PL/I VERSION.
6=C
7=C***RESULTS FOR THE NORMAL MODEL (PROBIT ANALYSIS) ARE ALSO PRODUCED BY
8=C  THIS ROUTINE. BOTH THE OMEGA MODEL (QUANTIT ANALYSIS) AND THE NORMAL
9=C  MODEL HAVE BEEN SCALED SO THAT THE PROBABILITY DENSITY FUNCTION F(X)
10=C  ATTAINS A MAXIMUM VALUE OF 1 AT THE ORIGIN (F(0) = 1).
11=C  I.E.,
12=C  THE OMEGA PDF IS:  $F(X) = 1 - W \cdot (V+1)$  WHERE W IS THE ABSOLUTE VALUE
13=C  OF (2*P-1).
14=C  THE NORMAL PDF IS:  $F(X) = \exp(-Y \cdot Y \cdot \pi)$  WHERE  $\pi = 3.14159265$ 
15=C
16=C**THE FOLLOWING STATEMENT DOUBLE PRECISIONS EVERY VARIABLE BEGIN NING
17=C  WITH THE LETTERS A THRU H AND O THRU Z.
18=  IMPLICIT DOUBLE PRECISION(A-H,O-Z)
19=  DIMENSION WHAT(7),XXX(50)
20=  COMMON /BLANK1/ XX(50),XN(50),XS(50),XP(50),XPHAT(50),XSAVE P(50)
21=  COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
22=  COMMON /BLANK3/ SAVEA(101),SAVEB(101)
23=  COMMON /BLANK4/ ISETS,ITERA,ISWANA,IVOP,ISTEP,KD,MN,IXSW
24=  COMMON /BLANK5/ SAVEV(50),V123(3),SAVELN(50),XL123(3)
25=  COMMON /COM1/ XK,XD,XI
26=  COMMON /COM2/ PI,SQ2PI,A1,A2,A4,XLIKE
27=  COMMON /ALPHAT/ VFIN,TITLE(9)
28=  COMMON /PANDV/ XPVAL(50),XVVAL(7)
29=  CHARACTER XCNTL*5,XTITLE*5,XEDVAL*5,XDOSES*5,XVPM*5,XFIN IS*5,
30=  1XBLANK*8,FINAL1*8,VFIN*5,TITLE*8,XLABEL*5
31=  DATA XCNTL,XTITLE,XEDVAL,XDOSES,XVPM,XFINIS /'CNTRL','TITLE',
32=  1'EDVAL','DOSES','VPM','FINIS'/
33=  DATA FINAL1/ '(FINAL)'/
34=  DATA XBLANK /' '/
35=  PI = 3.14159265359
36=  SQ2PI = DSQRT(2.*PI)
37=  ISETS = 0
38=C** READ CNTRL CARD
39=  1 READ(5,2,END=500) XLABEL,KD,LOGT,XLOGA,MN,IVOP,INOV,IPRTV
40=  2 FORMAT(A5,I2,I1,F3.0,I1,I1,I1,I1)
41=  ISETS = ISETS + 1
42=  NODOS = 0
43=  IERR = 0
44=  WRITE(7,3) ISETS
45=  3 FORMAT(1H1,'DATA SET ',I3)
46=  IF (XLABEL.EQ.XCNTL) GO TO 4
47=  WRITE(7,5)
48=  5 FORMAT(1H0,'* * * ERROR, CNTRL CARD IS NOT PRESENT OR IS NOT THE F
49=  1IRST CARD IN THIS DATA SET. PROGRAM IS TERMINATED')
50=  GO TO 500
51=  4 CONTINUE
52=  IF(LOGT.NE.3) GO TO 6
53=C**CHECK FOR VALID LOG TRANSFORMATION

```

```

54= IF(XLOGA.GT.1.0) GO TO 6
55= WRITE(7,7) XLOGA
56= 7 FORMAT(1H0,'* * * ERROR -- A= 'F8.3,' IS INVALID BASE FOR LOG TRA
57= INFORMATION')
58= 6 CONTINUE
59= IXSW = 1
60= ISWANA = 0
61= IEDCRD = 0
62= IVCARD = 0
63= VFIN = XBLANK
64=C**READ TITLE CARD
65= READ(5,8,END=500) XLABEL,(TITLE(I),I=1,9)
66= 8 FORMAT(A5,9A8)
67= IF(XLABEL.EQ.XTITLE) GO TO 9
68= WRITE(7,10) XLABEL,(TITLE(I),I=1,9)
69= 10 FORMAT(1H0,'* * * ERROR: TITLE CARD NOT ENCOUNTERED. THE FO LLOWING
70= 1CARD WAS READ IN: ',/1X,A5,9A8)
71= GO TO 500
72= 9 WRITE(7,11) (TITLE(I),I=1,9)
73= 11 FORMAT(1H0,9A8)
74=C**READ NEXT CARD
75= 50 READ(5,12) XLABEL,(WHAT(I),I=1,7)
76= 12 FORMAT(A5,5X,7F10.0)
77=C**CHECK IF EDVAL CARD
78= IF(XLABEL.NE.XEDVAL) GO TO 13
79= IEDCRD = 1
80= IF( (MN.GE.1).AND.(MN.LE.7)) GO TO 14
81= WRITE(7,15) MN
82= 15 FORMAT(1H0,'* * *ERROR: DATA FROM EDVAL CARD CANNOT BE RETR IEVED.
83= 1 CNTRL CARD INDICATES THAT THERE ARE 'I4,'VALUES')
84= IF(IERR.EQ.0) IERR = 1
85= GO TO 1000
86= 14 DO 16 I=1,MN
87= 16 XPVAL(I) = WHAT(I)
88= DO 17 I= 1,MN
89= IF(XPVAL(I).GT.0.).AND.(XPVAL(I).LT.1.) GO TO 17
90= WRITE(7,18) XPVAL(I)
91= 18 FORMAT(1H0,'* * *ERROR: THE ED VALUE OF P= 'D12.5,' IS OUT OF RAN
92= 1GE. MUST BE BETWEEN 0 AND 1')
93= GO TO 1000
94= 17 CONTINUE
95=C**CHECK IF VPARM CARD
96= 13 IF(XLABEL.NE.XVPARM) GO TO 19
97= IVCARD = 1
98= IF( (INOV.GE.1).AND.(INOV.LE.7)) GO TO 20
99= WRITE(7,21) INOV
100= 21 FORMAT(1H0,'* * *ERROR: DATA FROM VPARM CARD CANNOT BE RETR IEVED.
101= 1COLUMN 13 OF CNTRL CARD INDICATES THERE ARE 'I5,'VALUES')
102= IF(IERR.EQ.0) IERR = 1
103= GO TO 1000
104= 20 DO 22 I=1,INOV
105= XVVAL(I) = WHAT(I)

```

```

106= IF( (XVVAL(I).GT.(-1.0)).AND.(XVVAL(I).LE.20.))GO TO 23
107= WRITE(7,24) XVVAL(I)
108= 24 FORMAT(1H0,'* * * ERROR: INVALID VALUE OF V= 'D12.5,' WAS READ FR
109= 10M VPARM CARD. MUST BE GREATER THAN -1 AND LESS OR EQUAL TO 20.')
110= GO TO 1000
111=C**ONLY FIRST 3 DECIMAL PLACES OF V ARE USED.
112=C**THE FOLLOWING 3 STATEMENTS ARE IDENTICAL TO FLOOR(X) IN PL/I.
113=C THAT IS, THE LARGEST INTEGER .LE. TO X. THE FORTRAN FUNCTIO N
114=C IDINT(X) IS IDENTICAL TO FLOOR(X) FOR X.GT.0 , BUT NOT FOR
115=C NEGATIVE VALUES OF X.
116= 23 CONTINUE
117= VDEL = XVVAL(I)
118= DELSGN = DSIGN(0.5D0,VDEL)
119= V = DBLE( IDINT(VDEL*1000. + DELSGN) )/1000.
120= IF (XVVAL(I).LT.-.999) XVVAL(I) = -.999
121= 22 CONTINUE
122=C**CHECK IF DATA (I.E. 'DOSES') CARD HAS BEEN READ.
123= 19 IF (XLABEL.NE.XDOSES) GO TO 25
124= NODOS = NODOS + 1
125= IF (NODOS.LE.KD) GO TO 26
126= WRITE(7,27) KD
127= 27 FORMAT(1H0,'* * *ERROR: THE NO. OF DATA CARDS EXCEEDS THE V ALUE OF
128= 1 '13,' SPECIFIED IN COLUMNS 6-7 OF CNTRL CARD')
129= GO TO 1000
130= 26 XX(NODOS) = WHAT(1)
131= XS(NODOS) = WHAT(2)
132= XN(NODOS) = WHAT(3)
133= IF( (XS(NODOS).LT.0.0).OR.(XS(NODOS).GT.XN(NODOS)).OR.(XN(NODOS).L
134= 1E.0.0)) GO TO 29
135= GO TO 30
136= 29 WRITE(7,31) XX(NODOS),XS(NODOS),XN(NODOS)
137= 31 FORMAT(1H0,'* * *ERROR: ONE OR MORE INVALID DATA ITEMS: DOS E= 'D1
138= 12.5,' S= 'D12.5,' N= 'D12.5)
139= IF (IERR.EQ.0) IERR=1
140= GO TO 1000
141= 30 CONTINUE
142= 25 IF(XLABEL.NE.XFINIS) GO TO 50
143= IF( (MN.NE.0).AND.(IEDCRD.EQ.0))GO TO 32
144= GO TO 33
145= 32 WRITE(7,34) MN
146= 34 FORMAT(1H0,'* * *ERROR: MN= '15,' IN COL. 12 OF CNTRL CARD , BUT N
147= 10 EDVAL CARD IS PRESENT. DEFAULT EDVALUES WILL BE USED')
148= GO TO 35
149= 33 IF(MN.NE.0) GO TO 36
150=C**ASSIGN DEFAULT EDVALUES IF EDVAL CARD NOT PRESENT
151= 35 MN=7
152= XPVAL(1) =.01
153= XPVAL(2) =.05
154= XPVAL(3) =.10
155= XPVAL(4) =.50
156= XPVAL(5) =.90
157= XPVAL(6) =.95

```

```

158=   XPVAL(7) = .99
159=C**CHECK ON OPTION FOR V (COL. 13 ON CNTRL CARD)
160=C**CHECK FIRST FOR SEARCH PROCEDURE
161=   36 IF(IVOP.EQ.2) GO TO 37
162=   IF(IVOP.EQ.0) GO TO 41
163=C**CHECK IF VALUES OF V ARE READ IN (OPTION 2)
164=   IF( (IVOP.EQ.1).AND.(IVCARD.EQ.0) ) GO TO 38
165=   GO TO 39
166=   38 WRITE(7,40) IVOP
167=   40 FORMAT(1H0,'* * *ERROR: IVOP = ',I5,' IN COL. 13 OF CNTRL C ARD, BU
168=   1T NO VPARAM CARD IS PRESENT. DEFAULT VALUES ARE ASSIGNED')
169=   GO TO 41
170=   39 IF(IVOP.NE.0) GO TO 37
171=C**ASSIGN DEFAULT VALUES OF V.
172=   41 XVVAL(1) = -.9
173=   XVVAL(2) = -.5
174=   XVVAL(3) = 0.
175=   XVVAL(4) = 1.
176=   XVVAL(5) = 5.
177=   XVVAL(6) = 10.
178=   XVVAL(7) = 20.
179=   INOV = 7
180=   37 IF (NODOS.EQ.KD) GO TO 1000
181=   WRITE(7,43) KD,NODOS
182=   43 FORMAT(1H0,'* * *ERROR: ',I5,'DOSES ARE INDICATED ON CNTRL CARD(CO
183=   1L 6-7). ONLY ',I5,' DOSE CARDS WERE PRESENT')
184=   IF(IERR.EQ.0) IERR = 1
185=   GO TO 1000
186=C**CHECK IF ERROR. IF SO,GO TO NEW DATA SET.
187=   1000 IF(IERR.EQ.1) GO TO 1
188=   IF (LOGT.EQ.4) WRITE(7,44)
189=   44 FORMAT(1H0,'DOSAGE TRANSFORMATION: NONE')
190=   IF( (LOGT.LT.2).OR.(LOGT.GT.4) ) LOGT=1
191=   IF (LOGT.EQ.1) WRITE(7,45)
192=   45 FORMAT(1H0,'DOSAGE TRANSFORMATION: LOG(BASE 10)')
193=   IF (LOGT.EQ.2) WRITE(7,46)
194=   46 FORMAT(1H0,'DOSAGE TRANSFORMATION: NATURAL LN,BASE E')
195=   IF(LOGT.EQ.3) WRITE(7,47) XLOGA
196=   47 FORMAT(1H0,'DOSAGE TRANSFORMATION: LOG(BASE ',F5.0,' )')
197=   WRITE(7,48)
198=   48 FORMAT(1H0,17X,'TRANSFORMED NO. OF NUMBER')
199=   WRITE(7,49)
200=   49 FORMAT(1H ,7X,'DOSAGE',7X,'DOSE',8X,'SUBJECTS',4X,'RESPONDING',4X,
201=   1'PROPORTION')
202=   I=0
203=   162 I=I+1
204=   XSAVEP(I) = XS(I)/XN(I)
205=   XXI(I) = XX(I)
206=   XP(I) = XSAVEP(I)
207=   IF( (XS(I).EQ.0).OR.(XS(I).EQ.XN(I)) ) XP(I) = (XS(I)+ 0.5) /
208=   1(XN(I) + 1.0)
209=   IF(LOGT.EQ.4) GO TO 8001
210=   8000 IF ( LOGT.EQ.1) XXI(I) = DLOG10(XX(I))

```



```

211= IF(LOGT.EQ.2) XX(I) = DLOG(XX(I))
212= IF(LOGT.NE.3) XLOGA = 10.
213= IF(LOGT.NE.3) GO TO 8001
214= XTRANS = DLOG10(XLOGA)
215= XI(I) = DLOG10(XX(I))/XTRANS
216= 8001 CONTINUE
217= WRITE(7,63) XXX(I),XX(I),XN(I),XS(I),XSAVEP(I)
218= 63 FORMAT(1H0,3X,2(D12.5,2X),F6.0,7X,F6.0,9X,F9.6)
219= IF(I.LT.KD) GO TO 162
220= WRITE(7,51) KD
221= 51 FORMAT(1H0//,' NO. OF DOSAGE LEVELS = ',I4)
222= WRITE(7,52)
223= 52 FORMAT(1H0,'THE OPTION FOR V = ')
224= IF(IVOP.NE.2) GO TO 53
225= WRITE(7,54)
226= 54 FORMAT(1H+,20X,'SEARCH PROCEDURE')
227= IF(IPRTV.EQ.1) GO TO 55
228= WRITE(7,56)
229= 56 FORMAT(1H+,37X,'(PRINT COMPLETE RESULTS FOR V=0,1, AND FINA L V)')
230= GO TO 60
231= 55 WRITE(7,57)
232= 57 FORMAT(1H+,37X,'(PRINT COMPLETE RESULTS FOR ALL VALUES)')
233= GO TO 60
234= 53 CONTINUE
235= IF(IVOP.EQ.1) GO TO 58
236= WRITE(7,59) (XVVAL(I),I=1,INOV)
237= 59 FORMAT(1H+,20X,' DEFAULT VALUES: V= ', 7(F10.3) )
238= GO TO 60
239= 58 WRITE(7,61) (XVVAL(I),I=1,INOV)
240= 61 FORMAT(1H+,20X,' INPUTTED VALUES: V= ', 7(F10.3) )
241= 60 CALL MLEAB
242= CALL PRINT
243= ISWANA = 1
244= IF(IVOP.EQ.2) GO TO 65
245=C**PRODUCE RESULTS FOR FIXED VALUES OF V (IVOP = 1 OR 2)
246= DO 66 I=1,INOV
247= V= XVVAL(I)
248= CALL VRAT
249= CALL MLEAB
250= CALL PRINT
251= 66 CONTINUE
252= WRITE(7,3000) ISETS
253= 3000 FORMAT(1H0,'* * * E N D O F D A T A S E T',I3,'* * *')
254=C**GO TO NEW SET OF DATA
255= GO TO 1
256=C**SEARCH PROCEDURE: FIND V IN (-1 < V <= 20) THAT MAXIMIZES THE LIKELI
257= 65 ISTEP = 0
258= IVDONE = 3
259= DO 67 I=1,3
260= ISTEP =ISTEP + 1
261= TEMPV = DBLE(I) - 1.
262= V = TEMPV

```

```

263=  V123(I) = TEMPV
264=  SAVEV(I) = TEMPV
265=  CALL VRAT
266=  CALL MLEAB
267=  XL123(I) = XLIKE
268=  SAVELN(I) = XLIKE
269=  IF ( (I.LT.3).OR.(IPRTV.EQ.1) ) CALL PRINT
270=  IF ( (I.EQ.3).AND.(IPRTV.EQ.0))GO TO 68 271=      GO TO 67
272=  68 WRITE(7,70)
273=  70 FORMAT(1H1,58X,'SEARCH PROCEDURE',//1X,12X,'STEP',12X,'V',1 2X,'LN
274=  1L')
275=  WRITE(7,77) ISTEP,V,XLIKE
276=  67 CONTINUE
277=  IF( (XL123(2).GT.XL123(1)).AND.(XL123(2).GT.XL123(3)) ) GO TO 72
278=  GO TO 73
279=  72 DEL = -0.4
280=  DO 74 I = 1,2
281=  DEL =DEL/2.
282=  DO 75 J=1,2
283=  ISTEP = ISTEP + 1
284=  IVDONE = IVDONE + 1
285=  VDEL = V123(2) + DEL
286=  DELSGN = DSIGN(0.5D0,VDEL)
287=  V = DBLE( IDINT(VDEL*1000. + DELSGN) )/1000.
288=  V123(2*J-1) = V
289=  SAVEV(IVDONE) = V
290=  CALL VRAT
291=  CALL MLEAB
292=  XL123(2*J-1) = XLIKE
293=  SAVELN(IVDONE) = XLIKE
294=  IF(IPRTV.EQ.1) GO TO 175
295=  WRITE(7,77) ISTEP,V,XLIKE
296=  77 FORMAT(1H0,11X,I4,10X,F7.3,4X,D15.7)
297=C**FORMAT 77 IS F1 IN PL/I
298=  GO TO 176
299=  175 CALL PRINT
300=  176 CONTINUE
301=  DEL = -DEL
302=  75 CONTINUE
303=C*GO TO HILOW
304=  IF ( (XL123(3).GT.XL123(2)).OR.(XL123(1).GT.XL123(2)) ) GO TO 78
305=  74 CONTINUE
306=  V = 1.
307=C**GO TO V1
308=  GO TO 80
309=  73 CONTINUE
310=C**THIS IS HILOW
311=  78 IF (XL123(3).LE.XL123(2)) GO TO 81
312=  IF(IVDONE.NE.3) GO TO 82
313=  DEL = 2.
314=  GO TO 83
315=  82 DEL = -DEL
316=  83 XL123(1) = XL123(3)

```

```

317= V123(1) = V123(3)
318= GO TO 84
319= 81 IF (IVDONE.EQ.3) DEL = -.2
320=C**DELV
321= 84 CONTINUE
322= VDEL = V123(1) + DEL
323= DELSGN = DSIGN(0.5D0,VDEL)
324= V = DBLE( IDINT(VDEL*1000. + DELSGN) )/1000.
325= V123(2)= V
326= ISTEP = ISTEP + 1
327= DO 85 I = 1,IVDONE
328= IF(V.NE.SAVEV(I)) GO TO 85
329= IF(IPRTV.EQ.0) GO TO 86
330= WRITE(7,87) ISTEP,SAVEV(I)
331= 87 FORMAT(1H0,/,1X,'STEP',I3,': V= ',F5.2,' (PREVIOUSLY CALCULATED)')
332= 1)
333=C**FORMAT 87 IS F2 IN PL/I
334= GO TO 88
335= 86 WRITE(7,77) ISTEP,SAVEV(I),SAVELN(I)
336= 88 XL123(2) = SAVELN(I)
337=C**GO TO CHKL
338= GO TO 89
339= 85 CONTINUE
340= IVDONE = IVDONE + 1
341= SAVEV(IVDONE) = V
342= CALL VRAT
343= CALL MLEAB
344= IF(IPRTV.EQ.1) GO TO 91
345= WRITE(7,77) ISTEP,V,XLIKE
346= GO TO 92
347= 91 CALL PRINT
348= 92 XL123(2) = XLIKE
349= SAVELN(IVDONE) = XLIKE
350=C**CHKL
351= 89 IF(XL123(2).GT.XL123(1)) GO TO 93
352= DEL = -DEL/2.
353= XABS = DABS(DEL)
354= IF( (XABS.GE.0.5).AND.(V.GT.5.0) ) GO TO 94
355= IF( (XABS.GE.0.25).AND.(V.GT.2.0).AND.(V.LE.5.0) ) GO TO 94
356= IF( (XABS.GE.0.1).AND.(V.LE.2.0) ) GO TO 94
357= GO TO 95
358= 94 XL123(1) = XL123(2)
359= V123(1) = V123(2)
360=C**GO TO DELV
361= GO TO 84
362= 95 CONTINUE
363= GO TO 96
364= 93 CONTINUE
365=C**IF V=-.8 AND DEL=-.2, SET DEL = -.1
366= IF( DABS(V+.8D0).GT.1.0D-04) GO TO 8002
367= IF( DABS(DEL+.02D0).LE.1.0D-04 ) DEL = -.1D0
368= 8002 CONTINUE

```



```

422= WRITE(7,10) ISETS,(TITLE(I),I=1,9)
423= 10 FORMAT(1H0,'DATA SET ',I3,' ',9A8 )
424= WRITE(7,11) AINIT,BINIT
425= 11 FORMAT(1H0,/,1X,'INITIAL ALPHA = ',D15.7,7X,'INITIAL BETA = ',D15.
426= 17)
427= WRITE(7,12)
428= 12 FORMAT(1H0,'ITERATION',12X,'ALPHA',14X,'BETA')
429= DO 13 I=1,ITERA
430= WRITE(7,14) I,SAVEA(I),SAVEB(I)
431= 14 FORMAT(1H ,2X,I3,11X,D15.7,3X,D15.7)
432= 13 CONTINUE
433= AA2 = -A2
434= WRITE(7,15) A4,AA2,AA2,A1
435= 15 FORMAT(1H0,/,1X,'THE VARIANCE-COVARIANCE MATRIX FOR ALPHA AN D BETA:
436= 1',/,2(/2(6X,D15.7)) )
437=C**N_LINE
438= 16 ITENS = ITENS +9
439= NUM = MIN0(KD,ITENS+8)
440= WRITE(7,17) (XPHAT(I),I=ITENS,NUM)
441= 17 FORMAT(1H0,/,1X,'MLE FOR P: ',9(3X,F10.7) )
442= WRITE(7,18)(XSAVEP(I),I=ITENS,NUM)
443= 18 FORMAT(1H ,OBSERVED P: ',9(3X,F10.7) )
444= DO 19 I=1,KD
445= 19 PDIF(I) = XPHAT(I) - XSAVEP(I)
446= WRITE(7,20) ( PDIF(I),I=ITENS,NUM)
447= 20 FORMAT(1H ,DIFFERENCE : ',9(3X,F10.7) )
448= IF(NUM.LT.KD) GO TO 16
449= BHATSQ = BHAT*BHAT
450= XFIRST = A4 / BHATSQ
451= XLAST = -A2*2.0 / BHATSQ
452= XSEC = A1 / BHATSQ
453= IF (ISWANA.EQ.0) GO TO 21
454=C**QUANTIT EDVALUES
455= DO 22 I=1,MN
456= CALL QUANT1 (XPVAL(I),EDP(I) )
457= 22 EDP(I) = (EDP(I) - AHAT) / BHAT
458= GO TO 23
459=C**PROBIT RESUTLS
460= 21 CALL INVNOR(XPVAL,EDP,MN)
461= DO 24 I=1,MN
462= 24 EDP(I) = (EDP(I)/ SQ2PI - AHAT) / BHAT
463= 23 DO 25 I= 1,MN
464= VAREDP(I) = XFIRST + XSEC*EDP(I)*EDP(I) + EDP(I)*XLAST
465= 25 SEEDP(I) = DSQRT ( VAREDP(I) )
466= WRITE(7,26)
467= 26 FORMAT(1H0,/,6X,'P',9X,'ED ESTIMATE',10X,'VARIANCE',10X,'STD . ERROR
468= 1')
469= DO 27 I=1,MN
470= 27 WRITE(7,28) XPVAL(I),EDP(I),VAREDP(I),SEEDP(I)
471= 28 FORMAT(1H0,F10.7,3X,D15.7,2(4X,D15.7) )
472= DO 29 I=1,KD
473= PDIF(I) = PDIF(I)*PDIF(I)

```

```

474=   XYZ   = XPHAT(I)*(1. - XPHAT(I) )
475= 29 CHISQ = CHISQ + XN(I) * PDIF(I) / XYZ
476=   WRITE(7,30) XLIKE,CHISQ
477= 30 FORMAT(1H0,/,5X,'LN L = ',D15.7, 5X,'CHI SQUARE = ',D15.7 )
478=   RETURN
479=   END
480=   SUBROUTINE QUANCD(XPHAT,XX,F)
481=C**THIS ROUTINE CALCULATES THE CDF FOR THE OMEGA DISTRIBUTION. I. E.,
482=C GIVEN X, FIND P.
483=   IMPLICIT DOUBLE PRECISION (A-H,O-Z)
484=   COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
485=   COMMON /BLANK4/ ISETS,ITERA,ISWANA,IVOP,ISTEP,KD,MN,IXSW
486=   DIMENSION XPHAT(50),XX(50),F(50)
487=   DO 1 I= 1,KD
488=     Z = AHAT + BHAT * XX(I)
489=     KK = 0
490=     Q2 = 1.
491=     DO 2 J = 1,101
492=       PC =XPHAT(I)
493=C**ANOT
494=   3 IF (PC.LT.0.5) PC = 1.0 - PC
495=     CALL QUANT1(PC,HP)
496=     G = HP - DABS(Z)
497=     Q1 = 1. - (DABS(2.* PC - 1.)) ** (V+1.)
498=     XPHAT(I) = PC - G*Q1
499=     IF (XPHAT(I).LT.1.0) GO TO 4
500=C**FIND NEW INITIAL ESTIMATE OF P. IF INITIAL IS GREATER THE FINA L
501=C ESTIMATE ( FOR P GREATER THAN .5) , THEN CONVERGENCE IS GUARA NTEED.
502=     KK = KK + 1
503=     IF (KK.GT.1) GO TO 5
504=     PC = .9999D0
505=C**GO TO ANOT AND TRY AGAIN
506=     GO TO 3
507=   5 IF (KK.GT.2) GO TO 6
508=     PC = .99999999D0
509=     GO TO 3
510=C**NOTE THAT CONVERGENCE CRITERIA FOR P IS 0.00001 . HENCE IF T HE
511=C PROGRAM REACHES THIS POINT, SET P = .99999999
512=   6 XPHAT(I) = .99999999D0
513=C**GO TO FIN
514=     GO TO 10
515=   4 Q2 = DABS( XPHAT(I) - PC )
516=     IF(Q2.LT.0.00001) GO TO 10
517=     IF (J.LE.100) GO TO 2
518=     WRITE(7,11) I
519=   11 FORMAT(1H0,'NOTE: MORE THAN 100 ITERATIONS ARE REQUIRED FOR P',
520=   113,') IN SUBROUTINE QUANCD')
521=     GO TO 10
522=   2 CONTINUE
523=C**FIN
524=   10 IF(Z.LT.0.0) XPHAT(I) = 1. - XPHAT(I)
525=     IF(I) = 1. - (DABS(2.*XPHAT(I) - 1.)) ** (V+1.)
526=     RETURN

```

```

527= END
528= SUBROUTINE NORMCD (XPHAT,XX,KD,F)
529=C**THIS ROUTINE CALCULATES THE CDF FOR THE NORMAL DISTRIBUTION. I.E.,
530=C GIVEN X, FIND P.
531= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
532= COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
533= COMMON /COM2/ PI,SQ2PI,A1,A2,A4,XLIKE
534= DIMENSION XPHAT(50),XX(50),F(50)
535= DIMENSION B(5)
536= B0 = 0.2316419
537= B(1) = 0.31938153
538= B(2) = -0.356563782
539= B(3) = 1.781477937
540= B(4) = -1.821255978
541= B(5) = 1.330274429
542= DO 1 I=1,KD
543= CDF = 0.
544= Z = (AHAT + BHAT *XX(I) ) * SQ2PI
545= T = 1.0/ (1.0 + B0*DABS(Z) )
546= DO 2 J = 1,5
547= 2 CDF = CDF + B(J)*T**DBLE(J)
548= CDF =(CDF / SQ2PI) * DEXP(-Z*Z/2.)
549= XPHAT(I) = 1. - CDF
550= IF(Z.LT.0.0) XPHAT(I) = CDF
551= Z = AHAT + BHAT*XX(I)
552= 1 F(I) = DEXP(-Z*Z*PI)
553= RETURN
554= END
555= SUBROUTINE INVNR (P,Y,N)
556=C**THIS SUBROUTINE CALCULATES THE NORMAL DEViate (I.E.,MODIFIED P ROBIT)
557=C OF P. GIVEN P, FIND Y.
558= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
559= DIMENSION P(50),Y(50),CC(2),DD(3)
560= CC(1) = 0.802853
561= CC(2) = 0.010328
562= DD(1) = 1.432788
563= DD(2) = 0.189269
564= DD(3) = 0.001308
565= DO 1 I = 1,N
566= XNUM = 2.515517
567= XDEN = 1.0
568= PP = P(I)
569= IF(P(I).GT.0.5) PP = 1. - P(I)
570= T= DLOG(1.0/(PP*PP) )
571= T= DSQRT(T)
572= DO 2 J = 1,2
573= 2 XNUM = XNUM + CC(J)*T**DBLE(J)
574= DO 3 J= 1,3
575= 3 XDEN = XDEN + DD(J) * T**DBLE(J)
576= ZP = T - XNUM/XDEN
577= Y(I) = - ZP
578= IF (P(I).GT.0.5) Y(I) = ZP

```

```

579= 1 CONTINUE
580= RETURN
581= END
582= SUBROUTINE VRAT
583=C**THIS ROUTINE EXPRESSES V+1 AS A RATIONAL NUMBER. THIS IS NECES SARY
584=C FOR CALCULATION OF THE QUANTIT (IN SUBROUTINE QUANT1)
585= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
586= COMMON /COM1/ XK,XD,XI
587= COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
588= DIMENSION A(2,2)
589=C**FIND RATIONAL NUMBER F=V+1 = XK/XD
590= F= V+1
591= W1 = DBLE(IDINT(F) )
592= X= F - W1
593= X= DBLE( IDINT(X*1000. + .5) )
594= Y= 1000.
595= A(1,1) = W1
596= A(1,2)= 1.
597= A(2,1)= 1.
598= A(2,2)= 0.
599=C**IF X=0 , GO TO CALC
600= IF(X.EQ.0.) GO TO 2
601= B = DBLE( IDINT (Y/X) )
602= CALL MULT(A,B)
603= INUM = Y
604= DENOM = X
605= DO 1 I = 1,200
606= WORK1 = INUM
607= INUM = DENOM
608= DENOM = DMOD(WORK1,DENOM)
609= IF (DENOM.EQ.0.) GO TO 2
610= B= DBLE( IDINT(XNUM/DENOM) )
611= 1 CALL MULT(A,B)
612=C**CALC
613= 2 XK = A(1,1)
614= XD = A(2,1)
615= XI = -1
616= TWO = 2.
617= IF(DMOD(XK,TWO).EQ.0.) XI = 1
618= RETURN
619= END
620= SUBROUTINE MULT(A,B)
621= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
622= DIMENSION A(2,2),C(2,2)
623= C(1,1) = A(1,1)*B + A(1,2)
624= C(1,2) = A(1,1)
625= C(2,1) = A(2,1)*B + A(2,2)
626= C(2,2) = A(2,1)
627= DO 1 I=1,2
628= DO 1 J=1,2
629= 1 A(I,J) = C(I,J)
630= RETURN
631= END

```



```

632= SUBROUTINE QUANT1 (P,HP)
633=C**THIS ROUTINE CALCULATES THE QUANTIT OF P. I.E.,GIVEN P,FIND H( P).
634= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
635= COMMON /COM1/ XK,XD,XI
636= COMMON /COM2/ PI,SQ2PI,A1,A2,A4,XLIKE
637= COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
638= F=V+1.
639= IF (P.NE.0.5) GO TO 1
640= HP = 0.
641= RETURN
642= 1 PP = DABS( 2.*P -1.)
643= PP1= DEXP( DLOG(PP)/XD)
644= PP2= DEXP( 2.*DLOG(PP)/XD)
645= SUM1 = 0.
646= SUM2 = 0.
647= SUM3 = 0.
648= IMAX = IDINT( (XK-1.)/2.)
649= IF (IMAX.LT.1) GO TO 200
650= DO 20 I=1,IMAX
651= TEMPI = DBLE(I)
652= WORK1 = DCOS(2.*PI*TEMPI*XD/XK)
653= WORK2 = DCOS(2.*PI*TEMPI/XD)
654= WORK3 = DSIN(2.*PI*TEMPI*XD/XK)
655= WORK4 = DSIN(2.*PI*TEMPI/XK)
656= SUM1 = SUM1 - (XD/XK) * WORK1 * DLOG(1.- 2.*PP1*WORK2 + PP2 )
657= WORK5 = PP1*WORK4/(1. - PP1*WORK2)
658= 20 SUM2 = SUM2 + 2.*(XD/XK)*WORK3 * DATAN(WORK5)
659= 200 CONTINUE
660= IMAX = IDINT( (XD-1.)/XD)
661= IF (IMAX.LT.1) GO TO 201
662= DO 30 I = 1,IMAX
663= TEMPI = -DBLE(I)
664= WORK1 = TEMPI * F + 1.
665= 30 SUM3 = SUM3 - DEXP( WORK1 * DLOG(PP) - DLOG(WORK1) )
666= 201 HP = SUM1 + SUM2 + SUM3 - (XD/XD)*DLOG(1.-PP1)
667= 1 + XD*(1.+XI)*DLOG(1. + PP1)/(2.*XD)
668= HP = HP/2.D0
669= IF (P.LT.0.5) HP = -HP
670= RETURN
671= END
672= SUBROUTINE MLEAB
673=C**THIS ROUTINE CALCULATES THE MLE'S FOR ALPHA,BETA
674= IMPLICIT DOUBLE PRECISION (A-H,O-Z)
675= COMMON /BLANK1/ XX(50),XN(50),XS(50),XP(50),XPHAT(50),XSAVE P(50)
676= COMMON /BLANK2/ AHAT,BHAT,AINIT,BINIT,V
677= COMMON /BLANK3/ SAVEA(101),SAVEB(101)
678= COMMON /BLANK4/ ISETS,ITERA,ISWANA,IVOP,ISTEP,XD,MN,IXSW
679= COMMON /BLANK5/ SAVEV(50),V123(3),SAVELN(50),XL123(3)
680= COMMON /COM2/ PI,SQ2PI,A1,A2,A4,XLIKE
681= DIMENSION Y(50),F(50)
682= EPSI = 0.001
683= MAXIT = 100

```

```

684=  XLIKE = 0.
685=  ITERA = 0
686=  IF(ISWANA.EQ.0) GO TO 1
687=  DO 2 I=1,KD
688=  CALL QUANT1(XP(I),Y(I))
689=  2 XPHAT(I) = XP(I)
690=  GO TO 3
691=  1 CALL INVNOR(XP,Y,KD)
692=C**
693=  DO 20 I = 1,KD
694=  20 Y(I) = Y(I)/SQ2PI
695=C**INITIAL ALPHA,BETA BY LEAST SQUARES
696=  3 IF(IXSW.EQ.0) GO TO 4
697=  SUMX = 0.
698=  XTX = 0.
699=  DO 5 I = 1,KD
700=  SUMX = SUMX + XX(I)
701=  5 XTX = XTX + XX(I)*XX(I)
702=  SSX = XTX - (SUMX*SUMX)/DBLE(KD)
703=  XBAR = SUMX/DBLE(KD)
704=  IXSW = 0
705=  4 CONTINUE
706=  XTY = 0.
707=  SUMY = 0.
708=  DO 6 I = 1,KD
709=  XTY = XTY + XX(I)*Y(I)
710=  6 SUMY = SUMY + Y(I)
711=  YBAR = SUMY/DBLE(KD)
712=  SSP = XTY - (SUMX * SUMY)/DBLE(KD)
713=  BHAT = SSP/SSX
714=  AHAT = YBAR - BHAT*XBAR
715=C**NEWTON-RAPHSON PROCEDURE
716=C
717=C**ITER
718=  WRITE(7,51)
719=  51 FORMAT(1H0,'** 1 **')
720=  7 IF(ITERA.EQ.0) GO TO 8
721=  SAVEA(ITERA) = AHAT
722=  SAVEB(ITERA) = BHAT
723=  GO TO 9
724=  8 AINIT = AHAT
725=  BINIT = BHAT
726=  9 IF(ISWANA.EQ.0) GO TO 10
727=  CALL QUANCD(XPHAT,XX,F)
728=  GO TO 11
729=  10 CALL NORMCD(XPHAT,XX,KD,F)
730=  11 A1=0
731=  A2=0
732=  A4=0
733=  B1=0
734=  B2=0
735=  WRITE(7,52)

```

AD-A138 029

EXPLORATION OF DOSE-RESPONSE TECHNIQUES WITH SOME
APPLICATIONS TO A SIMUL. (U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH SCHOOL OF ENGI... L G KEHL
DEC 83 AFIT/MA/GOR/83D-4 F/G 9/2

2/2

UNCLASSIFIED

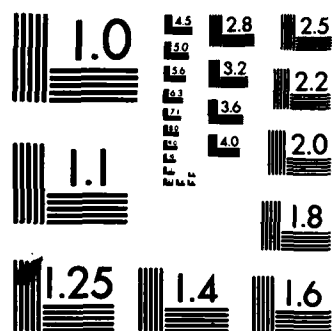
NL

END

FILED

DEC 83

AFIT



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

```

736= 52 FORMAT(1H0,'** 2 **')
737= DO 12 I = 1,KD
738= WRITE(7,100) I,XX(I),XPHAT(I)
739= 100 FORMAT(1H0,'I,XX(I),XPHAT(I)= ',I5,2(F10.5))
740= W = XN(I) / XPHAT(I)*(1. - XPHAT(I))
741= FF = F(I)*F(I)
742= A1 = A1 + W*FF
743= A2 = A2 + XX(I)*W*FF
744= A4 = A4 + XX(I)*XX(I)*W*FF
745= BB = W*F(I)*(XSAVEP(I) - XPHAT(I))
746= B1 = B1 + BB
747= 12 B2 = B2 + BB*XX(I)
748= WRITE(7,53)
749= 53 FORMAT(1H0,'** 3 **')
750= DET = A1*A4 - A2*A2
751= WRITE(7,54)
752= 54 FORMAT(1H0,'** 4 **')
753= DET = 1./DET
754= ADELTA = DET*(A4*B1 - A2*B2)
755= BDELTA = DET*(A1*B2 - A2*B1)
756= ABSAD = DABS(ADELTA)
757= ABSBD = DABS(BDELTA)
758= IF (DMAX1(ABSAD,ABSBD).GE.EPSI).AND.(ITERA.LT.MAXIT) GO TO 14
759= GO TO 15
760= 14 AHAT = AHAT + ADELTA
761= BHAT = BHAT + BDELTA
762= ITERA = ITERA + 1
763=C**GO TO ITER
764= GO TO 7
765= 15 IF (DMAX1(ABSAD,ABSBD).GE.EPSI).AND.(ITERA.EQ.MAXIT)WRITE (7,16)
766= 1MAXIT
767= 16 FORMAT(1H0,'THE ITERATION PROCESS HAS BEEN STOPPED. CONVERGENCE HAS
768= NOT BEEN ATTAINED AFTER ',I4,' ITERATIONS.')
769= A1 = DET*A1
770= A2 = DET*A2
771= A4 = DET*A4
772= XLIKE = 0.
773= DO 17 I=1,KD
774= YN = XN(I)
775= YS = XS(I)
776=C****Q1 = DLGAMA(YN + 1) - DLGAMA(YS+1) - DLGAMA(YN - YS + 1)
777=C** KEHL FIX DLGAMA NOT ON IMSL
778= XKEHL1=DABS(GAMMA(YN+1))
779= XKEHL2=DABS(GAMMA(YS+1))
780= XKEHL3=DABS(GAMMA(YN-YS+1))
781= Q1=DLOG(XKEHL1)-DLOG(XKEHL2)-DLOG(XKEHL3)
782= 17 XLIKE = XLIKE + Q1 + YS*DLOG(XPHAT(I)) + (YN-YS)*DLOG(1.-XPHAT(I))
783= RETURN
784= END

```

APPENDIX F

Computer Programs for Section VI

The first two programs, SYMTST and ASYMTST, calculate the test statistic in equations 6.3 and 6.7 respectively. The input data should be on TAPE 5 in the following order:

- 1.) First card - the logit regression coefficients B_0 and B_1 .
- 2.) Second card - title up 10 characters.
- 3.) Third card - N , the number of dose levels.
- 4.) Fourth through $N+3$ cards - the dose, number of subjects, number of responses

The test statistic and its variance are both printed on the interactive system display and written to TAPE 7.

The next two programs, SYMFIT and ASYMFT, uses an incremental procedure to fit the data to the models given by equations 6.1 and 6.5 respectively. The value of λ is incremented and new values of the least squares coefficients are calculated until the log likelihood function is maximized. The input to the programs should be on TAPE 5 in exactly the same order as 2 through 4 above, the first card is NOT used by these programs.

```

1=  PROGRAM SYMTST
2=  REAL N(20),X(20),R(20),THE(20),IBL(2),IBB(2,2),ILL,TCUBE(20 )
3=  +,E(20)
4=  CHARACTER NAME*10
5=  READ(5,*)B0,B1
6=  READ(5,('(A10)')NAME
7=  READ(5,*)NUM
8=  DO 10 I=1,NUM
9=  READ(5,*)X(I),N(I),R(I)
10=  X(I)=ALOG(X(I))
11=  IF(R(I).EQ.0) THEN
12=    R(I)=.5
13=  END IF
14=  TEMP=B0+B1*X(I)
15=  TCUBE(I)=TEMP**3
16=  THE(I)=1/(1.+EXP(-TEMP))
17=10  CONTINUE
18=  U=0.0
19=  ILL=0.0
20=  DO 20 I=1,2
21=    IBL(I)=0.0
22=  DO 30 J=1,2
23=    IBB(I,J)=0.0
24=30  CONTINUE
25=20  CONTINUE
26=  DO 40 I=1,NUM
27=    FACT=N(I)*(THE(I)*(1.-THE(I)))
28=    ILL=ILL+(FACT*TCUBE(I)**2)/144
29=    PART=((R(I)-N(I)*THE(I))*TCUBE(I))/12
30=    U=U+PART
31=    IBL(1)=IBL(1)+(FACT*TCUBE(I))/12
32=    IBL(2)=IBL(2)+(FACT*X(I)*TCUBE(I))/12
33=    IBB(1,1)=IBB(1,1)+FACT
34=    IBB(1,2)=IBB(1,2)+FACT*X(I)
35=    IBB(2,2)=IBB(2,2)+(FACT*X(I)**2)
36=40  CONTINUE
37=  DET=IBB(1,1)*IBB(2,2)-IBB(1,2)**2
38=C  REDEFINE IBB TO BE IBB INVERSE
39=  TEMP=IBB(1,1)/DET
40=  IBB(1,1)=IBB(2,2)/DET
41=  IBB(2,2)=TEMP
42=  IBB(1,2)=-IBB(1,2)/DET
43=  IBB(2,1)=IBB(1,2)
44=C  FINISH INVERSE ROUTINE
45=  T1=0.0
46=  T2=0.0
47=  DO 50 I=1,2
48=    T1=IBL(I)*IBB(1,1)+T1
49=    T2=IBL(I)*IBB(1,2)+T2
50=50  CONTINUE
51=  VAR=ILL-(T1*IBL(1)+T2*(IBL(2)))
52=  STDV=SQRT(VAR)
53=  PRINT*,U,STDV

```

54= TESTV=U/STDV
55= PRINT*,TESTV
56= WRITE(7,'(A10,3F20.10)')NAME,U,STDV,TESTV
57= END


```

1=  PROGRAM ASYMTS
2=  REAL N(20),X(20),R(20),THE(20),IBL(2),IBB(2,2),ILL,FACT(20)
3=  +,E(20)
4=  CHARACTER NAME*10
5=  READ(5,*)B0,B1
6=  READ(5,*(A10)')NAME
7=  READ(5,*)NUM
8=  DO 10 I=1,NUM
9=  READ(5,*)X(I),N(I),R(I)
10=  X(I)=ALOG(X(I))
11=  IF(R(I).EQ.0) THEN
12=    R(I)=.5
13=  END IF
14=  TEMP=B0+B1*X(I)
15=  E(I)=EXP(TEMP)
16=  THE(I)=E(I)/(1.+E(I))
17=10  CONTINUE
18=  U=0.0
19=  ILL=0.0
20=  DO 20 I=1,2
21=    IBL(I)=0.0
22=  DO 30 J=1,2
23=    IBB(I,J)=0.0
24=30  CONTINUE
25=20  CONTINUE
26=  DO 40 I=1,NUM
27=    FACT(I)=THE(I)+ALOG(1.-THE(I))
28=    ILL=ILL+(FACT(I)**2)*N(I)/E(I)
29=    U=U+(FACT(I)/THE(I))*(R(I)-N(I)*THE(I))
30=    IBL(1)=IBL(1)+FACT(I)*N(I)*(1.-THE(I))
31=    IBL(2)=IBL(2)+FACT(I)*N(I)*X(I)*(1.-THE(I))
32=    IBB(1,1)=IBB(1,1)+N(I)*THE(I)*(1.-THE(I))
33=    IBB(1,2)=IBB(1,2)+N(I)*X(I)*THE(I)*(1.-THE(I))
34=    IBB(2,2)=IBB(2,2)+N(I)*(X(I)**2)*THE(I)*(1.-THE(I))
35=40  CONTINUE
36=  DET=IBB(1,1)*IBB(2,2)-IBB(1,2)**2
37=C  REDEFINE IBB TO BE IBB INVERSE
38=  TEMP=IBB(1,1)/DET
39=  IBB(1,1)=IBB(2,2)/DET
40=  IBB(2,2)=TEMP
50=  IBB(1,2)=-IBB(1,2)/DET
51=  IBB(2,1)=IBB(1,2)
52=C  FINISH INVERSE ROUTINE
53=  T1=0.0
54=  T2=0.0
55=  DO 50 I=1,2
56=    T1=IBL(I)*IBB(1,1)+T1
57=    T2=IBL(I)*IBB(1,2)+T2
58=50  CONTINUE
59=  VAR=ILL-(T1*IBL(1)+T2*(IBL(2)))
60=  STDV=SQRT(VAR)
61=  PRINT*,U,STDV
62=  TESTV=U/STDV

```

63= PRINT*,TESTV
64= WRITE(7,'(A10,3F20.10)')NAME,U,STDV,TESTV
65= END

```

1=  PROGRAM SYMFIT
2=  IMPLICIT DOUBLE PRECISION(A-H,O-Z)
3=  COMMON/SET1/X(20),XP(20),TAU(20),XLAMDA,AHAT,BHAT,N(20)
4=  COMMON/SET2/NUM,IXSW
5=  REAL ISTAT,IDF,IQUE
6=  DOUBLE PRECISION MLEHAT,R(20),PFIN(20),THE(20)
7=  CHARACTER NAME*20
8=  IXSW=1
9=  CHECK=10000.
10=  CHKLIK=-999999999.
11=  PRINT*,' NAME?'
12=  READ(*,'(A20)')NAME
13=  READ(5,*)NUM
14=  IDF=NUM-2.
15=  DO 10 IN=1,NUM
16=    READ(5,*)X(IN),N(IN),R(IN)
17=    X(IN)=DLOG(X(IN))
18=    IF(R(IN).EQ. 0)R(IN)=.5
19=    XP(IN)=R(IN)/N(IN)
20=10  CONTINUE
21=  INCLAM=0
22=  DO 20 J=1,200
23=    XLAMDA=DBLE(INCLAM)/100.
24=    INCLAM=INCLAM+1
25=    XLLF=0.0
26=    CALL LSE
27=    STAT=0
28=    XLIKE=0.0
29=    DO 30 I=1,NUM
30=      FACT=.5*XLAMDA*(AHAT+BHAT*X(I))
31=      IF(DABS(FACT).LT. 1. .AND. FACT.NE. 0.) THEN
32=        T1=(1.+FACT)**(1/XLAMDA)
33=        T2=(1.-FACT)**(1/XLAMDA)
34=        THE(I)=T1/(T1+T2)
35=      ELSE IF(FACT.EQ. 0.0) THEN
36=        THE(I)=1/(1+EXP(-AHAT-BHAT*X(I)))
37=      ELSE IF(FACT.LE. -1.) THEN
38=        THE(I)=.9999999999
39=      ELSE
40=        THE(I)=.0000000001
41=      ENDIF
42=      EXPECT=N(I)*THE(I)
43=      DIFFSQ=(R(I)-EXPECT)**2
44=      TEST=DIFFSQ/(EXPECT*(1.-THE(I)))
45=      STAT=STAT+TEST
46=      XLIKE=XLIKE+R(I)*DLOG(THE(I))+N(I)-R(I))*DLOG(1.-THE(I))
47=30  CONTINUE
48=  IF(STAT.LT.CHECK)THEN
49=    DO 40 KK=1,NUM
50=      PFIN(KK)=THE(KK)
51=40  CONTINUE
52=  CHECK=STAT
53=  CHKLIK=XLIKE

```

```

54=     MLEHAT=XLAMDA
55=     B0=AHAT
56=     B1=BHAT
57=     END IF
58=20   CONTINUE
59=     WRITE(7,'(F6.4,2F20.15)'MLEHAT,B0,B1
60=     WRITE(7,*)CHKLIX
61=     WRITE(7,'(A20)'NAME
62=     WRITE(7,*)'THERE ARE ',NUM,' DATA POINTS.'
63=     WRITE(7,*)
64=     WRITE(7,*)
65=     WRITE(7,*)
66=,    WRITE(7,*)' MTBF  LAUNCHES  ABORTS   PRED ABT   PROB
151,
67=     DO 50 I=1,NUM
68=         EXPECT=N(I)*PFIN(I)
69=         DIFFSQ=(R(I)-EXPECT)**2
70=         TEST=DIFFSQ/(EXPECT*(1.-PFIN(I)))
71=         WRITE(7,'(3F7.0,7X,F9.5,5X,F6.5,3X,F7.4)'X(I),REAL(N(I)
),R(I)
72=     +,EXPECT,PFIN(I),TEST
73=50    CONTINUE
74=     ISTAT=REAL(CHECK)
75=     CALL MDCH(ISTAT,IDF,IQUE,IER)
76=     PCHI=1.-IQUE
77=     WRITE(7,*)
78=     WRITE(7,*)' TEST STATISTIC          CHI-SQ TAIL PROB'
79=     WRITE(7,'(3X,F10.6,15X,F7.5)'CHECK,PCHI
80=     WRITE(7,*)
81=     END
82=     SUBROUTINE LSE
83=     IMPLICIT DOUBLE PRECISION(A-H,O-Z)
84=     COMMON/SET1/X(20),XP(20),TAU(20),XLAMDA,AHAT,BHAT,N(20)
85=     COMMON/SET2/NUM,IXSW
86=     DOUBLE PRECISION W(20)
87=     DO 1 I=1,NUM
88=     IF(XLAMDA.EQ. 0.0) THEN
89=         TEMP=XP(I)/(1.-XP(I))
90=         TAU(I)=DLOG(TEMP)
91=     ELSE
92=         TEMP1=(1.-XP(I))*XLAMDA
93=         TEMP2=XP(I)*XLAMDA
94=         XNUM=2*(TEMP2-TEMP1)
95=         XDENOM=XLAMDA*(TEMP2+TEMP1)
96=         TAU(I)=XNUM/XDENOM
97=     END IF
98=1     CONTINUE
99=     IF(IXSW .EQ. 0) GOTO 4
100=     SUMX=0.
101=     SUMW=0.0
102=C   SET UNW=0 IF UNWEIGHTED LS IS WANTED SET TO
103=C   ANYTHING ELSE FOR WEIGHTED LS
104=     UNW=1.

```

```

105=   ITX=0.
106=   DO 5 I=1,NUM
107=       W(I)=DBLEQN(I))*XP(I)*(1.-XP(I))
108=       IF(UNW.LEQ.0)THEN
109=           W(I)=1.
110=           SUMW=SUMW+W(I)
111=       ELSE
112=           SUMW=SUMW+W(I)
113=       END IF
114=       SUMX=SUMX+X(I)*W(I)
115=       ITX=ITX+(X(I)**2)*W(I)
116=5   CONTINUE
117=   SSX=ITX-(SUMX**2)/SUMW
118=   XBAR=SUMX/SUMW
119=   IXSW=0
120=4   CONTINUE
121=   XTY=0.
122=   SUMY=0.
123=   DO 6 I=1,NUM
124=       XTY=XTY+X(I)*TAU(I)*W(I)
125=       SUMY=SUMY+TAU(I)*W(I)
126=6   CONTINUE
127=   YBAR=SUMY/SUMW
128=   SSP=XTY-(SUMX*SUMY)/SUMW
129=   BHAT=SSP/SSX
130=   AHAT=YBAR-BHAT*XBAR
131=   RETURN
132=   END

```

```

135= PROGRAM ASYMFT
136= IMPLICIT DOUBLE PRECISION(A-H,O-Z)

137= COMMON/SET1/X(20),XP(20),TAU(20),XLAMDA,AHAT,BHAT,N(20)
138= COMMON/SET2/NUM,IXSW
139= REAL ISTAT,IDF,IQUE
140= DOUBLE PRECISION MLEHAT,R(20),PFIN(20),THE(20)
141= CHARACTER NAME*60
142= IXSW=1
143= CHECK=10000.
144= CHKLIK=-9999999999
145= PRINT*,' NAME?'
146= READ(*,'(A60)')NAME
147= READ(5,*)NUM
148= IDF=NUM-2.
149= DO 10 IN=1,NUM
150=   READ(5,*)X(IN),N(IN),R(IN)
151=   X(IN)=DLOG(X(IN))
152=   IF(R(IN).EQ. 0.)R(IN)=.5
153=   XP(IN)=R(IN)/N(IN)
154=   IF(XP(IN).EQ.1 )XP(IN)=.9999
155=10 CONTINUE
156= INCLAM=-500
157= DO 20 J=1,1000
158=   XLAMDA=DBLE(INCLAM)/1.
159=   INCLAM=INCLAM+1
160=   XLLF=0.0
161=   CALL LSE
162=   STAT=0
163=   XLIKE=0.0
164=   DO 30 I=1,NUM
165=     FACT=XLAMDA*EXP(AHAT+BHAT*X(I))
166=     IF(FACT.GT. -1. .AND. FACT.NE. 0.) THEN
167=       T1=(1.+FACT)**(-1./XLAMDA)
168=       THE(I)=1.-T1
169=     ELSE IF(FACT.EQ. 0.0) THEN
170=       THE(I)=1.-EXP(-EXP(AHAT+BHAT*X(I)))
171=     ELSE IF(FACT.LE. -1.) THEN
172=       THE(I)=.9999999999
173=   ENDIF
174=   EXPECT=N(I)*THE(I)
175=   DIFFSQ=(R(I)-EXPECT)**2
176=   TEST=DIFFSQ/(EXPECT*(1.-THE(I)))
177=   STAT=STAT+TEST
178=   XLIKE=XLIKE+R(I)*DLOG(THE(I))+(N(I)-R(I))*DLOG(1.-THE(I))
179=30 CONTINUE
180= IF(XLIKE.GT.CHKLIK)THEN
181=   DO 40 KK=1,NUM
182=     PFINK(K)=THE(KK)
183=40 CONTINUE
184= CHECK=STAT
185= CHKLIK=XLIKE
186= MLEHAT=XLAMDA

```

```

187=      B0=AHAT
188=      B1=BHAT
189=      END IF
190=20    CONTINUE
191=      WRITE(7, '(F10.5,2F20.15)' )MLEHAT,B0,B1
192=      WRITE(7,*)CHKLIK
193=      PRINT*,MLEHAT
194=      WRITE(7, '(A60)' )NAME
195=      WRITE(7,*)'THERE ARE ',NUM,' DATA POINTS.'
196=      WRITE(7,*)
197=      WRITE(7,*)
198=      WRITE(7,*)
199=      WRITE(7,*)' MTBF  LAUNCHES  ABORTS   PRED ABT   PROB
199=      1ST
200=      DO 50 I=1,NUM
201=          EXPECT=N(I)*PFIN(I)
202=          DIFFSQ=(R(I)-EXPECT)**2
203=          TEST=DIFFSQ/(EXPECT*(1.-PFIN(I)))
204=      WRITE(7, '(3F7.0,7X,F9.5,5X,F6.5,3X,F7.4)  ',REAL(N(I)
204=      ),R(I))
205=      +,EXPECT,PFIN(I),TEST
206=50     CONTINUE
207=      ISTAT=REAL(CHECK)
208=      CALL MDCH(ISTAT,IDF,IQUE,IER)
209=      PCHI=1.-IQUE
210=      WRITE(7,*)
211=      WRITE(7,*)' TEST STATISTIC      CHI-SQ TAIL PROB'
212=      WRITE(7, '(3X,F10.4,15X,F7.5)' )CHECK,PCHI
213=      WRITE(7,*)
214=      END
215=      SUBROUTINE LSE
216=      IMPLICIT DOUBLE PRECISION(A-H,O-Z)
217=      COMMON/SET1/X(20),XP(20),TAU(20),XLAMDA,AHAT,BHAT,N(20)
218=      COMMON/SET2/NUM,IXSW
219=      DOUBLE PRECISION W(20)
220=      DO 1 I=1,NUM
221=          IF(XLAMDA.EQ. 0.0) THEN
222=              TEMP=1./(1.-XP(I))
223=              TAU(I)=DLOG(DLOG(TEMP))
224=          ELSE
225=              TEMP1=(1.-XP(I))*(-XLAMDA)
226=              TEMP2=(TEMP1-1.)/XLAMDA
227=              TAU(I)=DLOG(TEMP2)
228=          END IF
229=1      CONTINUE
230=      IF(IXSW .EQ. 0.) GOTO 4
231=      SUMX=0.
232=      SUMW=0.0
233=C     SET UNW=0 IF YOU WANT UNWEGHTED LS,IF NEED WEIGHTED
234=C     SET UNW TO ANYTHING BUT 1 (ONE)
235=      UNW=0.
236=      ITX=0.
237=      DO 5 I=1,NUM

```

```

238=      W(I)=DBLE(N(I))*XP(I)*(1.-XP(I))
239=      IF(UNW.EQ.0.)THEN
240=      W(I)=1.
241=      SUMW=SUMW+W(I)
242=      ELSE
243=      SUMW=SUMW+W(I)
244=      END IF
245=      SUMX=SUMX+X(I)*W(I)
246=      XTX=XTX+(X(I)**2)*W(I)
247=5     CONTINUE
248=      SSX=XTX-(SUMX**2)/SUMW
249=      XBAR=SUMX/SUMW
250=      IXSW=0
251=4     CONTINUE
252=      XTY=0.
253=      SUMY=0.
254=      DO 6 I=1,NUM
255=      XTY=XTY+X(I)*TAU(I)*W(I)
256=      SUMY=SUMY+TAU(I)*W(I)
257=6     CONTINUE
258=      YBAR=SUMY/SUMW
259=      SSP=XTY-(SUMX*SUMY)/SUMW
260=      BHAT=SSP/SSX
261=      AHAT=YBAR-BHAT*XBAR
262=      RETURN
263=      END

```


Bibliography

1. Aranda-Ordaz, F.J. "On Two Families of Transformations to Additivity for Binary Response Data," Biometrika, 68, 2: 357-363 (1981).
2. Armitage, P. and Doll, R. "Stochastic Models for Carcinogenesis," Proc. Forth Berkeley Symp. Math. Statist. Prob., 4: 19-38 (1961).
3. Barr, David R. Applications of Probit and Logit to the Prediction of the Probability of An Abort in the AEP Model. unpublished report. Department of Mathematics, Air Force Institute of Technology, WPAFB, Ohio. undated
4. Bartlett, M.S. "The Use of Transformations," Biometrics, 3: 39-52 (1947).
5. Berkson, Joseph "A Statistically Precise and Relatively Simple Method of Estimating the Bio-Assay with Quantal Response, Based on the Logistic Function," J. Amer. Stat. Assoc., 48: 565-599 (1953).
6. Box, G.E.P. and Cox, D.R. "An Analysis of Transformations," J. Royal Stat. Soc., Series B, 26: 211-252 (1964).
7. Brown, Charles C. "On a Goodness of Fit Test for The Logistic Model Based on Score Statistics," Comm. in Stat. Theory and Meth., 11, 10: 1087-1105 (1982).
8. Chambers, E.A. and Cox, D.R. "Discrimination Between Alternative Binary Response Models," Biometrics, 53, 3 and 4: 573-578 (1967).
9. Crump, K.S. "Dose-Response Problems in Carcinogenesis," Biometrics, 35: 157-167 (March 1979).
10. Crump, K.S., Guess, H.A. and Deal, K.L. "Confidence Intervals and Test of Hypotheses Concerning Dose-Response Relations Inferred from Animal Carcinogenicity Data," Biometrics, 33: 437-451 (September 1977).
11. Copenhaver, T.W. and Mielke, P.W. "Quantal Analysis: A Quantal Assay Refinement," Biometrics, 33: 175-186 (March 1977).
12. Cox, D.R. Analysis of Binary Data. London: Methuen & Co. Ltd., 1970.
13. Daffer, P.Z., Crump, K.S. and Masterson, M.D. "Asymptotic Theory for Analyzing Dose-Response Survival Data with Applications to the Low-Dose Extrapolation Problem," Math. Biosciences, 50: 207-230 (1980).

14. DeGroot, Morris H. Probability and Statistics. Reading: Addison-Wesley Pub. Co., 1975.
15. Finney, D.J. Probit Analysis. Cambridge: The University Press, 1947.
16. Finney, D.J. Statistical Method in Biological Assay. London: Griffen, 1964.
17. Fishbein, Lawrence "Overview of Some Aspects of Quantitative Risk Assessment," J. Toxicology and Environmental Health, 6: 1275-1296 (Sept/Nov 1980)
18. Goldfield, S.M., Quandt, R.E. and Trotter, H.F. "Maximization by Quadratic Hill-Climbing," Econometrica, 34: 541-551 (July 1966).
19. Guess, H.A. and Crump, K.S. "Low-Dose-Rate Extrapolation of Data from Animal Carcinogenicity Experiments--Analysis of a New Statistical Technique," Math. Biosciences, 32: 15-36 (1976).
20. Guess, H.A. and Crump, K.S. "Maximum Likelihood Estimation of Dose-Response Functions Subject to Absolutely Monotonic Constraints," The Annals of Stat., 6, 1: 101-111 (1978).
21. Guess, H.A., Crump, K.S. and Peto, R. "Uncertainty Estimates for Low-Dose-rate Extrapolation of Animal Carcinogenicity Data," Cancer Research, 37: 3475-3483 (October 1977).
22. Guerrero, U.M. and Johnson, R.A. "Use of the Box-Cox Transformations with Binary Response Models," Biometrika, 69, 2: 309-314 (1982).
23. Harville, D.A. "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," J. Amer. Stat. Assoc., 72, 358: 320-339 (June 1977).
24. Hoel, D.G. "Animal Experimentation and its Relevance to Man," Envir. Health Perspectives, 32: 25-30 (1979).
25. Kale, B.K. "On the Solution of Likelihood Equations by Iteration Process. The Multiparametric Case," Biometrika, 49, 3 and 4: 479-486 (1962).
26. Kehl, L.G. Preliminary Analysis for PAVE PILLAR. unpublished report, Concepts and Evaluation Group, Air Force Avionics Laboratory, WPAFB, Ohio, April 1982.
27. Kennedy, W.J. and Gentle, J.E. Statistical Computing. New York: Marcel Dekker, Inc., 1980.
28. Magnus, A., Mielke, P.W., and Copenhaver, T.W. "Closed Expressions for the Sum of an Infinite Series with Application to Quantal Response Assays," Biometrics, 33: 221-223 (March 1977).

29. Neter, John and Maynes, E.S. "On the Appropriateness of the Correlation Coefficient with a 0,1 Dependent Variable," J. Amer. Stat. Assoc., 65, 330: 501-509 (June 1970).
30. Prentice, R.L. "A Generalization of the Probit and Logit Methods for Dose Response Curves," Biometrics, 32: 761-768 (December 1976).
31. Prentice, R.L. "Discrimination Among some Parametric Models," Biometrika, 62, 3: 607-614 (1975).
32. Raff, M.S. "On Approximating the Point Binomial," J. Amer. Stat. Assoc., 51, 274: 293-303 (June 1956).
33. Rai, Kamta and Ryzin, J.V. "A Generalized Multihit Dose-Response model for Low-Dose Extrapolation," Biometrics, 37: 341-352 (June 1981).
34. Theil, Henri Principals of Econometrics. New York: John Wiley and Sons, Inc., 1971.
35. Weatherburn, C.E. A First Course in Mathematical Statistics. Cambridge: The University Press, 1946.
36. Wetherill, G.B., Chen, H. and Vasudeva, R.B. "Sequential Estimation of Quantal Response Curves: A New Method of Estimation," Biometrika, 53, 3 and 4: 439-454 (1966).
37. Yates, F. "The Use of Transformations and Maximum Likelihood in the Analysis of Quantal Experiments Involving Two Treatments," Biometrika, 42: 382-403 (1955).

Vita

Lieutenant Larry G. Kehl was born in Johnstown, Pennsylvania on 12 April 1951. He enlisted in the United States Air Force in July 1969. For over ten years he was an aircraft avionics technician serving in the United States and, for five years, in the far east. While in Taiwan he married the former Li Hua Hong.

Lieutenant Kehl, while on active duty, received a Bachelor of Science degree in Mathematics from Northern Michigan University, he graduated Magna Cum Laude. He received his United States Air Force commission in April of 1980 from Officers Training School. His first assignment as a commissioned officer was to the Avionics Laboratory at WPAFB. In June of 1982 he entered the School of Engineering's Operations Research program at the Air Force Institute of Technology.

Permanent Address: 5829 Troy Villa Blvd
Dayton, Ohio

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD A138 029

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS										
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited										
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE													
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/MA/GOR/83D-4			5. MONITORING ORGANIZATION REPORT NUMBER(S)										
6a. NAME OF PERFORMING ORGANIZATION		6b. OFFICE SYMBOL (If applicable)		7a. NAME OF MONITORING ORGANIZATION									
6c. ADDRESS (City, State and ZIP Code) Air Force Institute of Technology Wright-Patterson AFB, Ohio 45433			7b. ADDRESS (City, State and ZIP Code)										
8a. NAME OF FUNDING/SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (If applicable)		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER									
8c. ADDRESS (City, State and ZIP Code)			10. SOURCE OF FUNDING NOS. <table border="1"><tr><td>PROGRAM ELEMENT NO.</td><td>PROJECT NO.</td><td>TASK NO.</td><td>WORK UNIT NO.</td></tr><tr><td></td><td></td><td></td><td></td></tr></table>		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.					
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.										
11. TITLE (Include Security Classification) See Box 19													
12. PERSONAL AUTHOR(S) Larry G. Kehl, B.S., 1Lt, USAF													
13a. TYPE OF REPORT MS Thesis		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Yr., Mo., Day) 1983 December									
				15. PAGE COUNT 113									
16. SUPPLEMENTARY NOTATION <div style="text-align: right;">Approved for public release; distribution unlimited. LYNN E. VOLAVEN Dean for Research and Professional Development Air Force Institute of Technology (AFIT) Wright-Patterson AFB, OH 45433 7 Feb 84</div>													
17. COSATI CODES <table border="1"><tr><td>FIELD</td><td>GROUP</td><td>SUB. GR.</td></tr><tr><td>12</td><td>01</td><td></td></tr><tr><td></td><td></td><td></td></tr></table>			FIELD	GROUP	SUB. GR.	12	01					18. SUBJECT TERMS (Continue on reverse if necessary; use block number) Curve Fitting, Quantal Analysis, Logit, Probit, Quantit Bio Statistics, Bio Assay, Dose Response Analysis	
FIELD	GROUP	SUB. GR.											
12	01												
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Title: Exploration of Dose-Response Techniques with some Applications to a Simulation Problem Thesis Chairman: Dr. David R. Barr													
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED										
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. David R. Barr		22b. TELEPHONE NUMBER (Include Area Code) 513-255-3098		22c. OFFICE SYMBOL AFIT/ENC									

The United States Air Force has, over the past decade or so, invested much time and money in computer simulations and models. At the most basic level almost all of these simulations are input-output type procedures; variables of interest are changed to determine the effect they have on some other factor. This process is virtually indistinguishable from dose-response problems in bio-assay, hence, are capable of being analyzed by the same methods used in bio-assay. The two most commonly used techniques are probit and logit, but there are many other available techniques. An alternative to performing numerous, and sometimes redundant, simulations is to use these techniques whenever possible.

Data from the Avionics Evaluation Program (AEP) were used as the basis for estimating the probability of aircraft abort, based on the mean-time-between-failure (MTBF) of various equipment items, using four quantal assay techniques. The fits obtained from these models were compared to the more popular probit and logit results previously obtained by Dr. David Barr.

END

FILMED

384

DTIC